THIRD SEMIANNUAL TECHNICAL REPORT

(15 December 1970 - 15 June 1971)

FOR THE PROJECT

"RESEARCH IN

STORE AND FORWARD COMPUTER NETWORKS"

Principal Investigator

and Project Manager:

HOWARD FRANK (516) 671-9580

ARPA Order No. 1523

Contractor: Network Analysis Corporation

Contract No. DAHC15-70-C-0120

Effective Date: 15 October 1969

Expiration Date: 15 October 1971

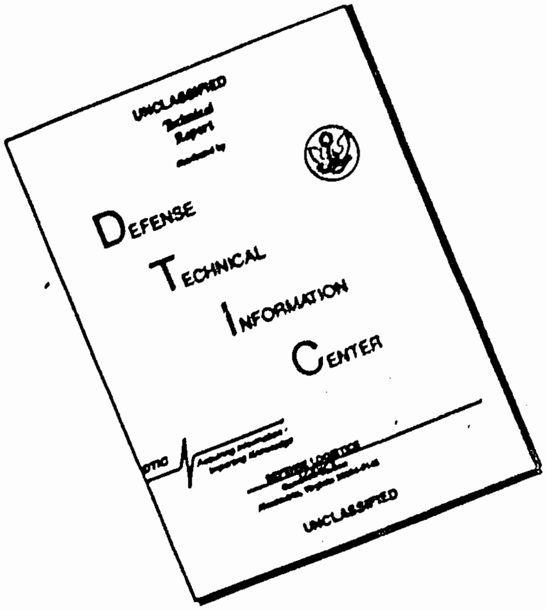
Amount of Contract: $276,107.00

Sponsored by

Advanced Research Projects Agency

Department of Defense

DISCLAIMER NOTICE



THIS DOCUMENT IS BEST QUALITY AVAILABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF PAGES WHICH DO NOT REPRODUCE LEGIBLY.

UNCLASSIFIED
Security Classification

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Network Analysis Corporation Beechwood, Old Tappan Road Glen Cove, New York 11542	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP None

3. REPORT TITLE

Semi-Annual Report No. 3
15 December 1970 - 15 June 1971

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Semi-Annual technical report. 15 December 1970 - 15 June 1971

5. AUTHOR(S) (Last name, first name, initial)

Network Analysis Corporation

6. REPORT DATE June 1971	7a. TOTAL NO. OF PAGES 115	7b. NO. OF REFS 11
8a. CONTRACT OR GRANT NO. DAHC 15-70-C-0120 b. PROJECT NO. c. ARPA Order No. 1523 d.	9a. ORIGINATOR'S REPORT NUMBER(S) ARPA Semi-Annual Report No. 3 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) — — —	

10. AVAILABILITY/LIMITATION NOTICES

This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES None	12. SPONSORING MILITARY ACTIVITY Advanced Research Projects Agency, Department of Defense

13. ABSTRACT This report considers the general problem of network reliability, the reliability analysis of the ARPA Computer Network, the relationship between cost and throughput in store-and-forward networks and the optimization of the growth of the ARPA Network. The approach to the reliability problem is to combine analytical and combinatorial results with simulation schemes to obtain several powerful procedures for reliability analysis. These procedures are then used to evaluate ARPA Network reliability. In the area of cost throughput trade-offs, the computer network optimization programs derive estimates for optimum network performance.

14. KEY WORDS

Computer networks, throughput, cost, reliability, survivability, ARPA Computer Network, store-and-forward

UNCLASSIFIED
Security Classification

SUMMARY

Technical Problem

The ARPA Contract with the Network Analysis Corporation involves the analysis and design of the ARPA Computer Network and the study and properties of networks of this type. During the reporting period the technical problems considered were the general problem of network reliability, the reliability analysis of the ARPA Network, the relationship between cost and throughput in store-and-forward systems, and the optimization of the growth of the ARPA network.

General Methodology

The approach to the reliability problem is to combine a number of analytical and combinatorial results with computer simulation schemes to derive several powerful procedures for reliability analysis. These procedures were then used to examine the ARPA Network reliability problem in detail. In the area of cost-throughput trade-off, the computer network optimization programs which have been under continuous development were used to derive estimates for optimum performance for specific network studies such as ARPA Network growth, design of large store-and-forward

networks, and the study of local distribution schemes.

Technical Results

A number of important technical results were derived during the reporting period:

1. Reliability analysis methods that are more than 1000 times more efficient than conventional schemes were developed.

2. The ARPA Network was shown to be highly reliable with respect to node and link failures.

3. Cost and throughput characteristics were determined for a 200-node store-and-forward network. These characteristics extend the results of our previous study of large networks which showed that such networks are economical to operate using the present technology of the ARPA Network.

Department of Defense Implications

The Defense Department has vital need of highly reliable and economical communications. The additional cost-throughput data substantiate our earlier conclusion (see Second Semiannual Report) that large computer communication networks can supply rapid and economical means for resource sharing and communications. The reliability studies provide the first step in guaranteeing that these networks will be highly reliable and survivable.

Implications for Further Research

The high efficiency of the new reliability analysis methods makes it possible for the first time to examine a number of analysis and design problems that were heretofore intractable. Research is now progressing to complement the economic studies of large systems with detailed studies of their reliability.

TABLE OF CONTENTS

1. GENERAL NETWORK RELIABILITY ANALYSIS

1.1. INTRODUCTION

We are concerned with a network in which the nodes and (undirected) links are in either a failed or an operative state. (We allow no intermediate states.) Two nodes can communicate if there is an alternating sequence of operative nodes and links starting with one of the two nodes and ending with the other such that each link in the sequence appears between its end nodes. We use as a measure of the networks inability to support communication the percent of node pairs which are not able to communicate. In Figure 1.1.1 the failed link (2, 5) and two failed nodes 1, and 5 are indicated by X's. The only pair which can communicate is 2,3.

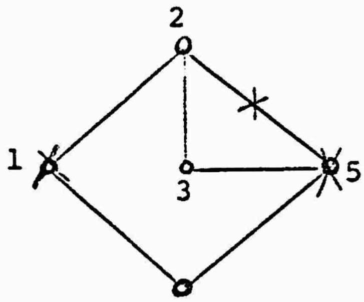


Figure 1.1.1

The number of node pairs in the net is (5) (4)/2 = 10. Nine of these do not communicate; thus, our measure of the network's

inadequacy is 9/10. It is sometimes convenient to use a simpler criterion. We say the network has <u>failed</u> if it is disconnected; i.e., if there exists any node pair which cannot communicate.

We now turn to the case where the links and nodes fail with some known probability. There are two essentially equivalent interpretations of the situation. The first case is when a "catastrophic" event occurs. For example, some natural disaster such as an earthquake or hurricane can cause the nodes and links to be knocked out with some probability. One can then ask what is the expected number of node pairs which can communicate if this event should occur. In the case of the second criterion, one would ask what is the probability the net will fail as a result of the event? The second interpretation is that the links or nodes fail and are repaired independently according to some stationary random process. We assume that the average time each node or link is operational is known and that it is equal to the probability of the node or link being in the operative state at any given time. In this case, the expected number of node pairs not communicating can be interpreted as the time average of pairs not communicating. The second criterion yields the percentage of the time the network is in a failed state.

If all the elements of the network which can fail have the same probability of failing, then combinatorial analysis methods can be useful. In the general case, a combination of analysis and simulation seems to be the most useful approach. We will first consider several special cases. In Section 2, we consider the case of infallible nodes and links with equal failure probabilities. This leads to strictly combinatorial analysis. In Section 3, we indicate our method for determining the components of a graph. This is not the usual method but turns out to be convenient for our application. In Section 4, we apply simulation in the case where no node failures are allowed but where the link failure probability can differ from link to link. The method described is particularly useful when the network is to be analyzed for a wide range of link failure probabilities. In Section 5, a second method is described which combines the combinatorial methods of Section 2 with simulation to analyze networks with invulnerable nodes. Sections 6 and 7 give the modifications required to the methods in Sections 4 and 5 respectively to apply them to networks with failing nodes and links.

1.2. COMBINATORIAL ANALYSIS OF NETWORKS WITH EQUAL LINK RELIABILITIES AND PERFECTLY RELIABLE NODES

The combinatorial analysis of networks with equal probabilities of link failures and perfectly reliable nodes is based on the work of Moore and Shannon [1956] on building reliable switching networks out of unreliable components. We will define the Moore-Shannon Function (MSF) to be the function h(p) that gives the probability of the net being disconnected as a function of the probability p of a link failing. The function a(p) = 1 - h(p), which gives the probability the net is connected, will be called the availability function for the net. Suppose there are NB links in the net and let q = 1 - p. Then there are $\binom{NB}{k}$ ways that exactly NB-k of the links can fail. Each of these events has the same probability $p^{NB-k}q^k$ of occurring. Thus, if we let C(k) be the number of ways exactly k remaining links can result in a disconnected net, we have

$$h(p) = \sum_{k=0}^{NB} C(k) \quad p^{NB-k} \; q^k \tag{1}$$

We also can define N(k) to be the number of ways k links can form a connected net. Clearly, $C(k) + N(k) = \binom{NB}{k}$. Thus, the availability problem "reduces" to the combinatorial problem of determining how many ways k links can result in a disconnected subnet.

A priori, we can also say something more about the form of (1). If the network has NN nodes, it takes at least NN-1 links to connect them. Thus

$$N(k) = 0. \qquad k=0,1, \ldots, NN-2$$

and

$$C(k) = \binom{NB}{k} \qquad k=0,1, \ldots, NN-2$$

Similarly, if we know the minimum number of links "c" which must be deleted to disconnect the network, we have

$$C(NB-k) = 0 \qquad k = 0, 1, \ldots, c-1.$$

If the network is initially disconnected $c = 0$.

In many applications c is quite small; for example, in the various versions of the ARPA Computer Network, $c = 2$. (See Chapter 2.)

Thus, there are only NB-NN-C+2 non-trivial terms in (1). This immediately gives us bounds on h(p).

$$\sum_{k=NB-NN+2}^{NB} \binom{NB}{k} p^k q^{NB-k} \leqq h(p) \leqq \sum_{k=c}^{NB} \binom{NB}{k} p^k q^{NB-k} \qquad (2)$$

These bounds are apparently due to Jacobs [1959]. If p is very close to 0, the last non-zero term in (1) determines the behavior

of $h(p)$. The last non-zero term is $C(NB-c)$. Similarly, if p is very close to 1, the first non-zero term in (1) dominates. This term $C(NN-1)$ is simply $\binom{NB}{NN-1}$ minus the number of trees in the network. $C(NB-c)$ can quite often be obtained easily by inspection and the number of trees can be obtained by the formula [Seshu and Reed; 1961, p. 157]:

$$\text{No. of trees} = \text{determinant}(UU^t)$$

where U is the reduced incidence matrix of the network.

We then get the approximation

$$\sum_{k=NB-NN+2}^{NB} \binom{NB}{k} p^k q^{NB-k} + C(NN-1)\, p^{NB-NN+1} q^{NN-1}$$

$$\leq h(p) \leq$$

$$C(NB-C)\, p^c q^{NB-c} + \sum_{k=c+1}^{NB} \binom{NB}{k} p^k q^{NB-k}.$$

The lower bound is sharp for p close to 1 and the upper bound for p close to 0. The bounds given in (3) can be further improved by using the fact that if a given subset of k links

disconnect the graph, any larger subset containing the first will also disconnect the graph. Similarly, if a subset of links is a connected subgraph of the network so is any subset containing it. This can be used to project lower bounds for C(k') given C(k) for k > k'. Similarly upper bounds can be obtained for C(k') given C(k) for k > k'.

One very general way to carry this out can be based on a powerful theorem by J. B. Kruskal [1963]. Kruskal defines an <u>abstract complex</u> to be a finite set of points together with a class of subsets with the subset closure property; that is, if a subset belongs to the class, then so do all its subsets. For any non-negative integer n he defines its <u>r-canonical representation</u> to be $(n_r, \ldots, n_i)$ where

$$n = \binom{n_r}{r} + \binom{n_{r-1}}{r-1} + \ldots + \binom{n_i}{i} \qquad (4)$$

and we first choose n_r as large as possible so that $\binom{n_r}{r} \leq n$ then we choose n_{r-1} as large as possible so that $\binom{n_r}{r} + \binom{n_{r-1}}{r-1} \leq n$ and so on until equality is achieved. Then for $r \leq r'$, he defines $f(n; r, r')$ to be the <u>greatest</u> number of r'-sets that occur in any complex having precisely n r-sets. If $r > r'$, he defines $f(n; r, r')$ to be the <u>smallest</u> number of r'-sets that occur in any complex having precisely n r-sets.

The result we are interested in is <u>Kruskal's theorem</u>

<u>Theorem 1.</u> If

$$n = \binom{N_r}{r} + \ldots \binom{N_i}{i} \text{ is canonical,}$$

then

$$f\ (n;\ r,\ r') = \binom{n_r}{r'} + \binom{n_{r-1}}{r'-1} + \ldots + \binom{n_i}{r'-r+i} \tag{5}$$

with the conventions that

$$\binom{0}{0} = 0,\ \binom{m}{k} = 0 \text{ for } m < 0 \text{ or } k < 0, \text{ or } m < k.$$

If a subnetwork with k links is disconnected, then a subnetwork with only k' (for k' < k) of the k links will also be disconnected.

Thus Kruskal's theorem gives us the following inequality:

$$C(k') \geqq f(C(k);\ k,\ k') \text{ for } k \geqq k'. \tag{6}$$

Similarly,

$$C(k') \leqq f(C(k);\ k,\ k') \text{ for } k \leqq k'. \tag{7}$$

Relations (6) and (7) are very useful. For each C(k) that we can calculate (or for that matter even get bounds on), we can get bounds on the remaining C(k'). In practice, one is usually concerned with a loosely-

connected net (c small) with a low probability of link failure. In such cases, one can exactly determine the first few of the terms $C(NB-c)$ $C(NB-c-1)$, ... (if necessary by enumeration), and then derive lower bounds for the remaining co-efficients using (6). This yields a very good lower estimate for $h(p)$ for p small. Good upper bounds on the coefficients are more difficult to determine. $C(NN-1)$, which is $\binom{NB}{NN-1}$ minus the number of trees, can be calculated by formula and upper bounds for the $C(k)$ for $k > NN-1$ arrived at by (7). Unfortunately, the terms which are most important for small p are the ones for which the estimates from (7) are least accurate.

The estimates for $C(k)$ implied by (6) and (7) are based purely on the fact that if a network is disconnected with one set S of operative links, it is also disconnected for any subset of S. The bound takes into account no other structure of the problem (not even the number of links as long as NB is sufficiently large). We now establish that while the bounds are sharp for complexes, they are not for network reliability problems. We will do this in a somewhat indirect way in order to also develop some machinery for later use.

A matroid $\langle S;\mathcal{F}\rangle$ [Fulkerson, 1968] is a finite set S and a non-empty family of subsets $\mathcal{F}$ of S satisfying:

(M1) No member of $\mathcal{F}$ is a proper subset of another.

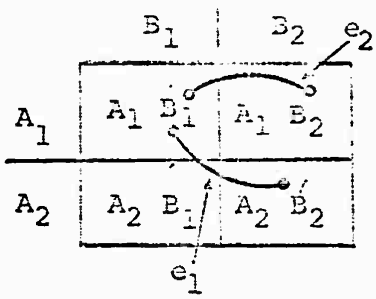
(M2) If $e_1 \neq e_2$, $C_1 \in \mathcal{F}$, $C_2 \in \mathcal{F}$, $e_1 \in C_1 \cap C_2$ and $e_2 \in C_1 - C_2$ then there exists $C_3 \in \mathcal{F}$ such that $e_2 \in C_3 \subset C(C_1 \cup C_2) - \{e_1\}$.

We are interested in matroids because of the following:

Theorem 2: Let $\mathcal{F}$ be the family of all minimal* link cut sets of a graph $\mathcal{G}$ with links S. $\langle S, \mathcal{F}\rangle$ is a matroid.

Proof: By definition, no member of $\mathcal{F}$ is a proper subset of another member. Suppose $C_1 \in \mathcal{F}$, $C_2 \in \mathcal{F}$ $e_1 \in C_1 \cap C_2$, $e_2 \in C_2 - C_1$.

We prove first that $C_1 \cup C_2 - \{e_1\}$ disconnects $\mathcal{G}$. C_1 partitions the nodes into two connected sets A_1 and A_2, and C_2 partitions the nodes into two connected sets B_1 and B_2. Suppose $e_1 = (m_1, n_1)$ (i.e., e_1 connects node m_1 to node n_1). Without loss of generality, we can assume $m_1 \in A_1$, $m_1 \in B_1$, $n_1 \in A_2$, $n_1 \in B_2$.



*A cut set is a set of links, which when deleted from a network leaves the resulting subnet disconnected. A minimal cut set is a cut set which contains no proper subset which is a cut set.

Partition S into $A_2 B_1 \cup A_1 B_2$ and $A_1 B_1 \cup A_2 B_2$ and let C'_3 be the set of links across the partition. $e_1 \notin C'_3$. Let $e' = (m', n') \in C'_3$. Suppose (without loss of generality) $m' \in A_2 B_1$; then if $n' \in A_1 B_1$, we have $e' \in C_1$. On the other hand, if $n' \in A_2 B_2$ then $e' \in C_2$. Thus $C'_3 \subset C_1 \cup C_2 - \{e_1\}$. Finally, we must show that there is a member C_3 of $\mathcal{J}$ such that $e_2 \in C_3 \subset C_3'$. Let C_3 be a subset obtained from C_3' by deleting from C_3' all links which leave the endpoints of e_2 in different components. There clearly is such a set since e_2 belongs to different components relative to the cut C_3'. The number of components remaining is exactly two. For if there were at least 3, say D_1, D_2, and D_3, then there must be two links in C_3 which connect D_1, D_2 and D_3 since the graph was originally connected. Only one of these two links can be e_2. Therefore, the other can be deleted from C_3 thus contradicting its minimality. This completes the proof of the theorem.

Suppose we consider (7) for $k = 2$, $k' = 3$ and $C(2) = 5$. Then (7) yields $C(3) \geq 2$. We now show that there exists no graph such that the number of connected graphs with 2 links is 5 and with 3 links is 2. More generally, we show that $C(k) \neq 2$ for all graphs with k and NB satisfying $\frac{1}{2}NB > k > 1$. This follows

directly from Theorem 2 and defining property M_2 of matroids.
Let $B = \{b_1, \ldots, b_k\}$ and $E = \{e_1, \ldots, e_k\}$ be non-identical sets of links which are disconnected. And let $\{\beta_1, \ldots, \beta_{NB-k}\}$ and $\{\epsilon_1, \ldots, \epsilon_{NB-k}\}$ be the links removed in each case. The total number of elements in $\beta = \{\beta_1, \ldots, \beta_{NB-k}\}$ and $\sum = \{\epsilon_1, \ldots, \epsilon_{NB-k}\}$ is $2NB-2k \geq NB$ since $NB \geq k$. Thus for some $i, \beta_i \in \sum$, and for some $j, \beta_j \notin \sum$ since $B \neq E$. If β and $\sum$ define minimal cuts, then there must exist a third minimal cut by the matroid property M2. But if $\sum$ is not minimal there exists $\sum' \subset \sum$ which is minimal. But in that case there would be more than two sets* S with $|S| = NB-k$ and $S \supset \sum'$.
On the other hand, consider the complex defined on the points 1, 2, 3, 4, 5, 6 with 3 sets (4, 5, 6) and (3, 5, 6) and 2 sets (5,6), (4,6), (4,5), (3,6), (3,5). These achieve the equality in (7).

This leads to the following interesting but yet unsolved problem:

Let $\langle S, \mathcal{F} \rangle$ be a matroid and let $\mathcal{F}'$ be the family of all subsets of S containing a set in $\mathcal{F}$. Given there are $C(k)$ sets in $\mathcal{F}'$ with cardinality k, what is a sharp lower bound on $C(k')$ for $k' > k$ and a sharp upper bound for $C(k')$ for $k' < k$.

---

* If X is a set, $|X|$ represents the number of elements in X.

Of course, it would be most helpful to be able to solve:

Let $\mathcal{C}$ be the cut sets of a finite graph with NB links and NN nodes. Given there are C(k) cuts with cardinality k, what is a sharp lower bound for C(k') for k' $\geq$ k and what is a sharp upper bound for C(k') for k' $\leq$ k.

It should be pointed out that Leggett [1968] gives similar and apparently stronger bounds to those mentioned here. However, his proof seems inadequate, and we have not been able to complete it.

A method of another sort due to Frank [1970] is based on the equivalent tree construction of Gomory and Hu [1961].

1.3. DETERMINING COMPONENTS OF NETWORKS

Consider a network $G = \langle \mathcal{N}, \mathcal{A} \rangle$ with node set $\mathcal{N}$ and link set $\mathcal{A}$. We wish to find the number of components of the network. Each node will be assigned a label indicating which component it is in. The algorithm is as follows:

<u>Step 0</u>: Start with $\mathcal{A}_0 = \emptyset$ and assign each node a separate label. Set $k = 0$. Go to Step 1.

<u>Step 1</u>: Add a link a_k to $\mathcal{A}_r$ to form $\mathcal{A}_{k+1}$. If $\mathcal{A}_{k+1} = \mathcal{A}$ or equivalently $k+1 = NB$, stop. Suppose $a_k = (m_k, n_k)$. Examine the labels of m_k and n_k. If they are the same, repeat Step 1 with* $k: = k+1$. If not, go to Step 2.

<u>Step 2</u>: Change all the node labels which are the same as the label of m_k (including m_k's label) to the label of n_k. Set $k: = k+1$ and go to Step 1.

When the algorithm terminates, each component is listed. It is important for future applications of the algorithm that we may introduce the links in Step 1 in any order we please.

It is convenient to maintain several other statistics of interest during the calculation. These might include the number of components, the number of nodes in each component, or the number of node pairs which are in the same component. This is

---

* The notation $k: = k+1$ means k is replaced by $k+1$.

carried out as follows. Initially, the number of components is NN, the number of pairs communicating NP is 0, and each component contains 1 node.

Each time we reach Step 2, we combine the two components, with say t_1 and t_2 nodes into a new component with $t_1 + t_2$ nodes. Also, we now have $t_1 \times t_2$ more node pairs which can communicate. Therefore, we set $NP = NP + t_1 \cdot t_2$. The number of components decreases by 1. Note that we can save computer time by terminating the algorithm wherever the network becomes connected; that is, when $NP = NN(NN-1)/2$ or equivalently the number of components is 1.

The next question we examine is how the addition or deletion of links affects the connectivity of $\mathcal{G}$. Assume we have completed the analysis for $\mathcal{G} = \langle \mathcal{N}, \mathcal{A} \rangle$ using the algorithm above. If we then add a link to $\mathcal{A}$ yielding the link set $\mathcal{A}'$, we need only repeat one cycle of the algorithm to complete the analysis of the new network. The simplicity of this operation is an important virtue of the algorithm. However, if we delete a link, a, from $\mathcal{A}$, the situation is more difficult. Suppose a = (m, n). Necessarily, m and n must be in the same component. We first label node m, we then label all nodes connected to m by a link in $\mathcal{A}' = \mathcal{A} - \{a\}$. In general, we label all nodes connected to a labeled node by a link in $\mathcal{A}'$. If n becomes labeled, the connectivity situation

is as before. If n cannot be labeled, let t_1 be the number of labeled nodes and t_o the number of nodes in the component containing a before it was removed. Then the number of components increases by one. The old component has t_1 elements, the new one $t_o - t_1$ and $NP = NP - t_1(t_o - t_1)$.

1.4. SIMULATION METHOD I: PERFECTLY RELIABLE NODES

The simulation method we are to describe here can be used to solve very general reliability problems. For simplicity, we first present it in its simplest context. In Section 1.6, we describe its generalizations.

Suppose we wish to apply simulation to the problem of determining the expected number of pairs communicating in a network (or to finding the probability of the network failing). Assume that links may fail with probability p but that the nodes do not fail. A direct scheme would be to choose a number randomly between 0 and 1 for each link. If the number is less than p, the link is removed; otherwise the link remains. Then we determine the number of node pairs communicating. We then choose another set of random numbers and compute the number of node pairs communicating and so on. The sample average obtained in this way would give us an estimate of the expected number of pairs communicating.

If we use the method described in Section 1.3. for computing the number of node pairs communicating, with little extra effort we can estimate the entire function of expected pairs communicating as a function of the probability of link failure. As before, we generate a random number between 0 and 1 for each link. We then

determine the largest random number associated with it and so on. Suppose the numbers in decreasing order are r_1, r_2, ..., r_{NB} for a given sample and let NP_1, NP_2, ..., NP_{NB} be the number of pairs communicating after the first link, the second link, ..., and the last link have been added. Then for $1 \geq p \geq r_1$ the sample value for this sample is 0; for $r_1 \geq p \geq r_2$ the sample value is NP_1, for $r_2 \geq p \geq r_3$ the sample value is NP_2 and so on. The entire procedure for determing the NP for all p is essentially the same as applying the algorithm once to the overall network. This very simple idea, which depends strongly on the form of the algorithm for determining the connectivity, saves considerable computation when the expected number of pairs communicating is desired for a range of link failure probabilities.

We now consider the case where the links have different probabilities of failing. Let us assume the nominal probability of link i failing is p_i. In our simulation scheme, we generate random numbers as before. The random number for link i is first divided by p_i, and then for each sample we sort the resulting numbers in decreasing order as before to obtain r_1, r_2, ..., r_{NB} and NP_1, NP_2, ..., NP_{NB}. This yields the sample mean NP as a function of p as before. However, p now has a different interpretation. For p = 1, the corresponding link probabilities are

$1 \cdot p_1$, $1 \cdot p_2$, ..., $1 \cdot p_{NB}$. For $p = \frac{1}{2}$, the expected number of pairs communicating is obtained for link failure probabilities, with one-half their nominal value. Thus, the method applied in the case of unequal link failure probabilities gives a sensitivity analysis where the probability of failing for each link is varied proportionally.

1.5. SIMULATION METHOD II: PERFECTLY RELIABLE NODES

The simulation technique to be described here is based on the combinatorial methods described in Section 1.2. As in that section, we partition the probability space into cases where no links fail, exactly one link fails, exactly two links fail, ..., all the links fail. The method depends on counting the number of ways the network can fail when k= 0, 1, ..., NB links fail. Instead of determining upper and lower bounds on the terms by combinatorial methods as we did in Section 1.2, we estimate the missing terms by sampling. This technique will be most useful in the situation of a loosely-connected network and where the probability of a link failing is small.

To introduce the ideas of this method, let us consider the network [Leggett, 1968] in Figure 1.5.1.

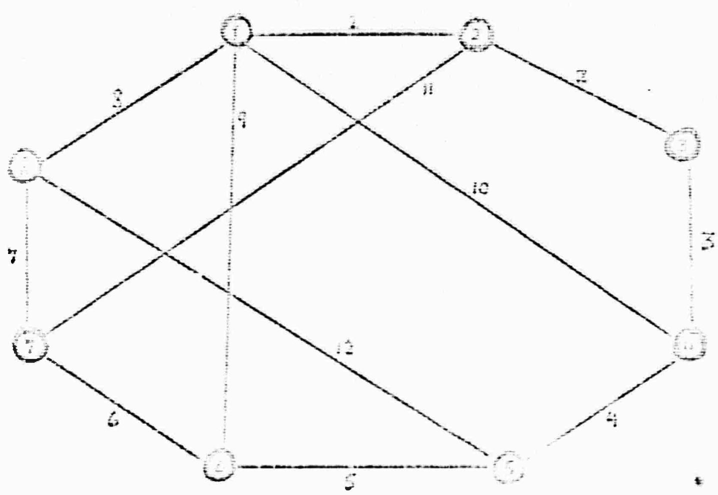


Figure 1.5.1.

In Table 1.5.1 we have tabulated for $k = 0, 1, \ldots, 12$, the number of failed nets with exactly k failed links ($= C(NB-k)$), the total number of nets with exactly k failed links ($= \binom{NB}{k}$), the contribution to the probability of net failure ($= C(NB-k)\, p^k q^{NB-k}$), and finally the probability of a net with exactly k failed links occurring ($= \binom{NB}{k}\, p^k q^{NB-k}$) for $p = .05$.

TABLE 1.5.1

Links Failed k	Nets Failed C(12-k)	Number of Nets $\binom{12}{k}$	$C(12-k)p^kq^{12-k}$	$\binom{12}{k}.p^rq^{12-k}$
0	0	1	.0	.5403
1	0	12	.0	.3412
2	1	66	.00149	.0987
3	18	220	.00141	.0178
4	124	495	.00051	.00205
5	465	792	.00009	.00017
6	924	924	.00001	.00001
7	792	792	small	small
8	495	495	.	.
9	220	220	.	.
10	66	66	.	.
11	12	12	.	.
12	1	1	small	small
Sum	3109	2^{12}	.00353	1.000

Given only the number of nodes, the number of links, and the size of the minimum cut set, we can calculate the values of all the C(k) except C(10), C(9), C(8), C(7) using the methods described in Section 1.2. Now suppose we were going to estimate these coefficients by sampling. The first thing we notice is that some of the coefficients have greater importance than others. For example, the probability of a network occurring with 10 operating links is .0987 while the probability of a network occurring with only 7 links is .00017, a difference by almost a factor of 600. This suggests that we expend "more effort" estimating the terms for the smaller number of link failures. In what follows, we make this notion more precise; additionally, we will define a systematic way for sampling which will take advantage of the form of our algorithm for finding the components of a graph.

A standard technique in such cases is found in the theory of stratified random sampling [Fisz, 1963] and is called proportional sampling. The strata in this case are those networks with no link failures, those networks with exactly one link failure, etc. Of these, we can eliminate from consideration those strata for which we already have complete information. In our numerical example we need only consider k = 2, 3, 4, 5;

that is, subnetworks with 7, 8, 9, or 10 links. Suppose we are allowed 100 samples. In proportional sampling, we sample each stratum a number of times proportional to the stratum's probability of occurrence. This suggests we sample: (100) (.0987)/ (.0987 + .0173 + .00205 + .00017) networks with two failed links; (100) (.0173)/(.0987 + ... + .00017) networks with 3 links failed; and so on. This becomes (after rounding to integers): 83 samples for networks with two link failures, 15 samples for networks with 3 link failures, 2 samples for networks with 4 link failures, and no samples for networks with 5 link failures. To make the statistical theory easier, we will assume we take one sample from nets with 5 links failed. Note that there are only 66 possible networks with 2 links failing so that instead of sampling 83 times with replacement among 66 networks, we clearly get better results by enumerating all 66 networks.

Since we have analyzed the network of Figure 1.5.1 in detail, we can compare the efficiencies of Simulation Method I and Simulation Method II. We define a random variable X so that $X = 1$ with probability equal to the probability of the network failing, P_F, and $X = 0$ otherwise. We then take n samples of X, say X_1, X_2, ..., X_n from nets whose links have failed randomly. Consider the variance

of $\overline{X} = (X_1 + \dots + X_n)/n$. This is $\sigma^2 = \frac{P_f(1-P_f)}{n}$. In Simulation Method I, we find the variance of $\overline{X}$ obtained by examining [illegible] 66 + 15 + 2 + 3 samples of the network. Thus

$$\sigma_I^2 = \frac{(.00353)\ (.99647)}{84} \cong 4.18 \times 10^{-5} .$$

In Simulation Method II we sample from the 220 nets with 3 links failed 15 times with replacement. Of the 220 nets, 18 of them are disconnected so P_f in this case is $\frac{18}{220}$ and the variance $\sigma_3^2 = \frac{1}{15}\left(\frac{18}{220}\right) = .0050$. For nets with 4 links failed, there are 495 possibilities of which 124 of them are disconnected. We take 2 samples in this case for a variance of $\sigma_4^2 = \frac{1}{2}\left(\frac{124}{495}\right)\left(\frac{371}{495}\right) = .0938$. Finally, we take one sample from the 792 networks with 5 failed links of which 456 are disconnected. $\sigma_5^2 = \left(\frac{456}{792}\right)\left(\frac{336}{792}\right) = .2749$.

The estimate for the probability of the net being disconnected is

$$h = \sum_{0}^{12} \overline{X}_k \binom{12}{k} p^k q^{12-k}$$

where $\overline{X}_k$ is the random variable which is the sample mean for the fraction of networks with k failed links which are disconnected. For all k except 3, 4, and 5, we know the fraction of failures exactly. For k=0, 1, all nets are connected; for k ≥ 6 all are disconnected. We enumerate all nets for k=2. Thus, the variance in each of these cases is 0. Remembering that the variance of

a constant times a random variable is equal to the constant squared times the variance of the random variable, we have

$$\sigma_{II}^{2} = [\binom{12}{3}\, p^{3}q^{12-3}]^{2}\, \sigma_{3}^{2} + [\binom{12}{4} p^{4}q^{12-4}]^{2}\, \sigma_{4}^{2}$$

$$+ [\binom{12}{5} p^{5}q^{12-5}]^{2}\, \sigma_{5}^{2}$$

$$= (.00141)^{2}(.005) + (.00051)^{2}\ (.094)$$

$$+ (.00009)^{2}\ (.274)$$

$$\cong 3.6 \times 10^{-8}$$

Thus, Simulation Method II is more efficient than Simulation Method I by a factor of more than 1000 as measured by the variance for this example. In general, this will be the case when p is close to 0, since when Simulation Method I is applied in such cases, most of the samples will be connected, and little information will be gained. However, in Simulation Method II since only the first few terms of the reliability function have any significance for the result, proportional sampling is very effective.

1.6. SIMULATION METHOD I: UNRELIABLE NODES AND LINKS

We first modify the method described in Section 1.3. for determining the components of a network. The modification is required to handle the situation where nodes can fail. Step 1 now becomes

Step 1': Add a link a_k to $\mathcal{A}_k$ to form $\mathcal{A}_{k+1}$. If $\mathcal{A}_{k+1} = \mathcal{A}$ or equivalently, $k+1 = NB$, stop. Otherwise, suppose $a_k = (m_k, n_k)$. Examine m_k and n_k and if either one is inoperative, or if they have the same labels, repeat Step 1' with $k := k+1$. If not, go to Step 2.

Similarly, modifications must be made to the procedures for adding or subtracting links. To add a node, one simply tries to add all the operative links incident to the node. To subtract a node, one deletes all the links incident to the node.

With the above modifications, Simulation Method I for node and link failures is very similar to the method without node failures. To begin, we make all the nodes and links inoperative and assign all the nodes to different components. We then generate NN + NB random numbers. The node or link corresponding to the largest of these is made operative and is added to the network. Then the node or link corresponding to the next largest random number is introduced and so on. The statistics are collected

in the same way as for the no node failure case. Unequal failure probabilities are also handled in the same way as before.

1.7. SIMULATION METHOD II: NODE AND LINK FAILURES

To adapt Simulation Method II to networks with node and link failures, we need only re-define the strata. In this case, they are defined by the number of link failing and the number of nodes failing. Thus, the first several strata are defined by 0 links failed - 0 nodes failed, 1 link failed - 0 nodes failed, 0 links failed - 1 node failed, 2 links failed - 0 nodes failed, 1 link failed - 1 node failed, 0 links failed - 2 nodes failed ... . If nodes have probability p_N of failing and links have probability p_A of failing, a network with exactly m failed links and n failed nodes has probability

$$\binom{NB}{m}\binom{NN}{n} p_N^{\,n} (1-p_N)^{NN-n} p_A^{\,m} (1-p_A)^{NB-m}$$

of occurring. Proportional random sampling can be used as before. We illustrate the procedure for the network defined by Figure 1.4.1. Let us suppose p_N = .05, and we wish to make approximately 100 samples. The probability of the first few strata are:

$$\binom{12}{0}\binom{8}{0}(.05)^0(.95)^8(.05)^0(.95)^{12} = .358$$

$$\binom{12}{1}\binom{8}{0}(.05)^1(.95)^7(.05)^0(.95)^{12} = .226$$

$$\binom{12}{0}\binom{8}{1}(.05)^0(.95)^8(.05)^1(.95)^{11} = .151$$

$$\binom{12}{2}\binom{8}{0}(.05)^0(.95)^8(.05)^2(.95)^{10} = .0655$$

$$\binom{12}{1}\binom{8}{1}(.05)^1(.95)^7(.05)^1(.95)^{11} = .0950$$

$$\binom{12}{0}\binom{8}{2}(.05)^2(.95)^6(.05)^0(.95)^{12} = .0282$$

As in the case where nodes do not fail, some of the strata can be analyzed explicitly in advance. These terms of course do not contribute to the variance of the estimate. For our example, the following strata are determined without sampling. If no nodes fail and less than two links fail, all pairs can communicate. If one node fails and no links fail, $\frac{(12)(11)}{2} = 12$ node pairs. communicate. Such networks account for .358+.226+.151 = .735 of the probability. There remains 1.-.735 = 265 to be accounted for. If we were allotted 1000 samples, and we intend to use proportional stratified random sampling, we would allocate

$1000\left(\frac{.0655}{.265}\right) \cong 247$ to the 66 nets with exactly 2 link failures,

$1000\left(\frac{.0950}{.265}\right) \cong 358$ to the 96 nets with one link failure and

one node failure and $1000\left(\frac{.0282}{.265}\right) \cong 106$ to the 28 nets with two node failures. These would all be enumerated leaving 1000-(66+96+28)=810 samples for the remaining strata. We would then consider strata corresponding to networks with exactly three elements (links or nodes) failing and so on as in Section 1.5.

2. RELIABILITY ANALYSIS OF THE ARPA NETWORK

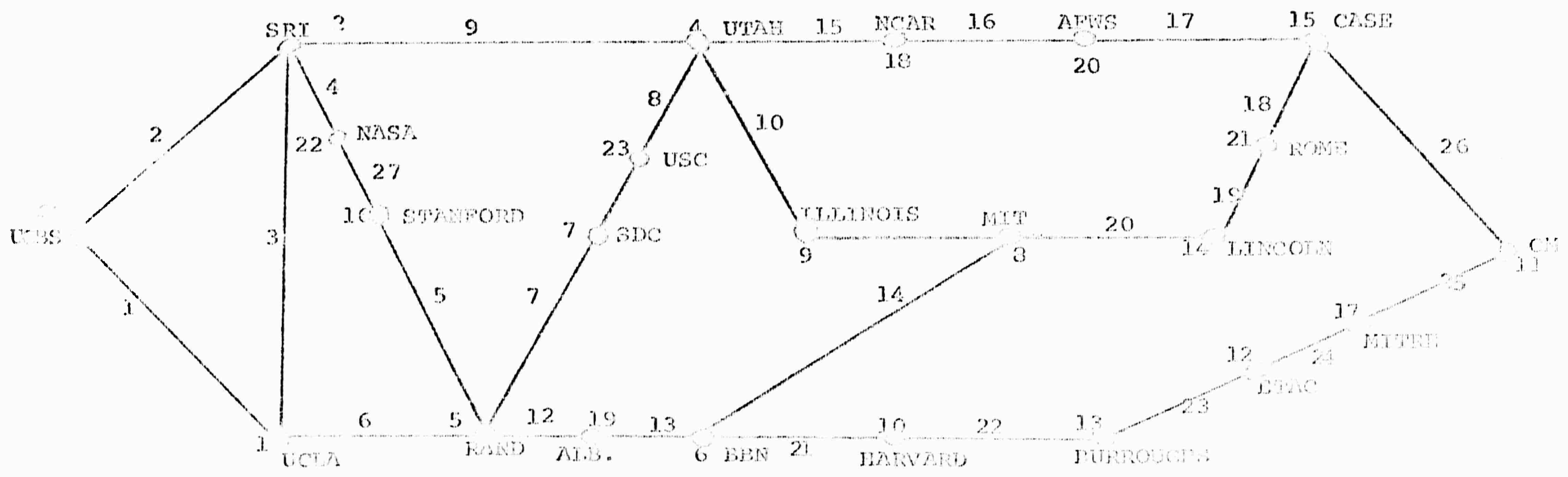
2.1. INTRODUCTION

The ARPA network is a store-and-forward computer network designed to interconnect many dissimilar computers located throughout the country. Each computer interfaces with the network by means of an Interface Message Processor (IMP). These IMPs are connected by fully duplex communication lines of typically 50 kilobit/sec. capacity. The reliability of the network and its availability to users is the subject of this chapter.

For analysis purposes, the ARPA network can be represented as a graph with lines corresponding to communication links and nodes corresponding to the Interface Message Processors. In our earlier work [Frank, Frisch, Chou, 1970] methods were described to choose network designs providing good response time at low cost. A minimum level of reliability was guaranteed by requiring that there exist at least two node disjoint paths between each pair of IMPs. Figure 2.1.1. represents a version of the ARPA network which is representative of the planned design at the end of 1971. This network consists of 23 nodes and 28 links and will be used throughout the chapter as an example.

In Section 2.2. we introduce two possible measures of reliability for the network. In Section 2.3. we analyze the ARPA

23 Node Net



Base Network for Reliability Analysis

FIGURE 2.3.1

network with respect to the first measure in the case where nodes are assumed to be invulnerable. In Section 2.4., we allow node and link failures while measuring reliability performance according to the second criterion. In addition, some modifications of the original network which increase the reliability of the network are explored.

2.2. TWO NETWORK RELIABILITY CRITERIA

Nodes and links can be in two states, failed or operative. Two nodes in the network can communicate if they both are operative and there exists an alternating sequence of operative nodes and links such that the first element is one of the two nodes, the last element in the sequence is the other node, and each link appears in the sequence between its end nodes. A simple and natural characterization of a failed or operative network is:

Criterion C1: A network is operative if every pair of operable nodes can communicate; otherwise, it is failed.

This is equivalent to saying that a network is operative if all the operating nodes are in one component; (a component is a maximal set of nodes in which each pair can communicate). Criterion C1 is not completely satisfactory because it does not indicate the "degree" of disruption a failed network has experienced.

For example, the failure of the single node 1 in the network shown in Figure 2.2.1. entirely prevents communication between the remaining nodes. In Figure 2.2.2., the failure of node 1 only prevents one operable node from communicating with the others.

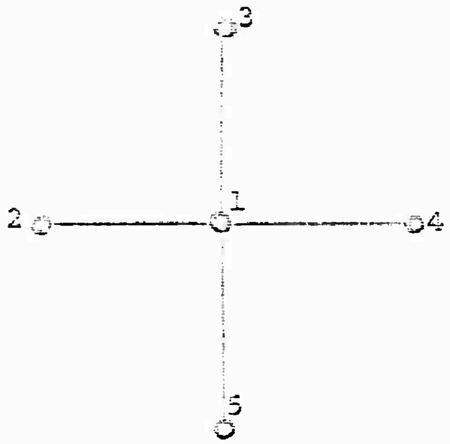


Figure 2.2.1.

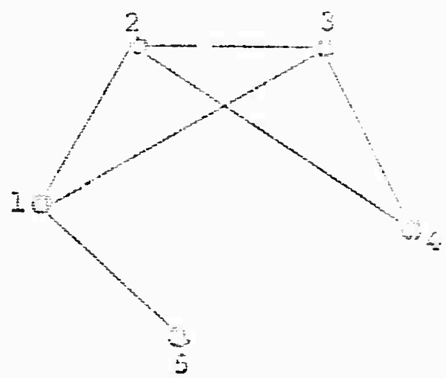


Figure 2.2.2.

<u>Criterion C2</u>: The relative connectedness of a network is the ratio of the distinct pairs which can communicate to the total number of distinct node pairs.

Suppose the nodes and links fail with known probabilities. We allow the possibility that each node and link can have a different probability of failing, but we insist that the links and nodes fail independently of one another. We seek to calculate in the case of Criterion C2 the expected number of node pairs which cannot communicate, and in the case of Criterion 1, the probability that the network is disconnected.

Our model serves to represent two situations. In the first, we think of some catastrophic event such as an earthquake, hurricane or the like. We assume that the nodes or links can be destroyed with some known probability and ask, for example, what is the expected number of node pairs which will be able to communicate after the event. The other situation is when links and nodes are continually failing and being repaired, and we wish to know either the time average of the number of node pairs communicating or the fraction of the time the network is connected. In this interpretation, the failure probability for each node and link is the percentage of the time it is not operational.

2.3. NETWORK CONNECTIVITY PROBABILITY

The IMPs in the ARPA system are very rugged, highly reliable units and preliminary information implies they will be much more reliable than the telephone lines which connect them. So as a first approximation, we can assume the nodes are perfectly reliable and look only at the effect of link failures. Moreover, to start with we will assume that all links have the same probability of failing. In this case, the problem can be analyzed combinatorially. Basically, if p represents the probability of a link failing, then the probability, $h(p)$, of the net failing can be written as

$$h(p) = \sum_{k=0}^{NB} C(k)\, p^{NB-k}\, q^{k} \qquad (1)$$

where $q = 1 - p$, NB is the number of l[illegible] in the net, and $C(k)$ is the number of subnets out of all possible subnets with k links which are disconnected. Thus, the original probabilistic problem reduces to the combinatorial problem of determining the $C(k)$. Several of the $C(k)$ can be specified immediately. If the net has NN nodes, it takes at least NN-1 links to connect the nodes together. Thus $C(k) = \binom{NB}{k}$ for $k < NN - 1$. On the other end, if it takes at least c links to disconnect the net (for the ARPA net, $c = 2$), $C(NB-k) = 0$ for $k = 0, 1, \ldots, c-1$. There are then NB-NN-c+2 of the $C(k)$ remaining undetermined. Since the ARPA net is designed to minimize cost, typically NB is not much larger than NN and $c = 2$. Thus, for the example in Figure 2.1.1., the number of non-trivial terms is 28-23-2+2 = 5. There are essentially 3 ways of obtaining the missing terms: enumerating of subnets with k links and counting the failed ones, sampling among subnets with k links, and estimating $C(k)$, or giving bounds on the missing terms. A more complete discussion of this is given in Chapter 1. The essential feature to keep in mind is in the case where p is very small (hopefully the situation for

the ARPA net) only networks with small numbers of simultaneous failures are likely to occur. In (1), therefore the non-trivial C(k) with the largest k are of most interest. Therefore, a reasonable line of attack is to enumerate the C(k) for the first few terms and estimate or sample for the remaining. We illustrate this procedure on the ARPA net of Figure 2.1.1. In Table 2.3.1. we enumerate the number of subnetworks with k= 0, 1, ..., 28 failed links and the number, in each case, of networks which are not connected when these numbers are known.

The remaining two terms can be accounted for in two ways. The first way takes advantage of the fact that if the removal of k links disconnects the net, then the removal of any k+1 links including the first k links will also disconnect the net. The details of the estimating procedure is given in Section 1.2. The resulting bounds on C(23) and C(24) are given in Table 2.3.2. The terms C(23) and C(24) can also be obtained by sampling. For small p, C(24) is much more crucial in determining the probability of the network failing. Thus, we expend more effort in determining C(24). The rationale we adopt is proportional stratified random sampling. We will assume we are interested in the range $0 \leq p \leq .1$. For p = .05, the probability of a [illegible]work having 23 operational links is $\binom{28}{23}(.05)^5(.95)^{23}$ = [illegible]

TABLE 2.3.1

EXACTLY KNOWN C(k) FOR 23 NODE 28 LINK ARPA NET

Number of links Operative	Number of links Failed	Number of Nets	Number of Failed Nets	Method of Determination
0	28	1	1	a
1	27	28	28	
2	26	378	378	
3	25	3276	3276	
4	24	20475	20475	
5	23	98280	98280	
6	22	376740	376740	
7	21	1184040	1184040	
8	20	3108105	3108105	
9	19	6906900	6906900	
10	18	13123110	13123110	
11	17	21474180	21474180	
12	16	30421755	30421755	
13	15	37442160	37442160	
14	14	40116600	40116600	
15	13	37442160	37442160	
16	12	30421755	30421755	
17	11	21474180	21474180	
18	10	13123110	13123110	
19	9	6906900	6906900	
20	8	3108105	3108105	
21	7	1184040	1184040	a
22	6	376740	349618	b
23	5	98280	?	
24	4	20475	?	
25	3	3276	827	c
26	2	378	30	c
27	1	28	0	d
28	0	1	0	d

Notes:
a: not enough links to connect 23 nodes
b: number of trees calculated by formula [Seshu & Reed, 1961, p. 157)
c: enumerated
d: less failed links than minimum cut set

TABLE 2.3.2

BOUNDS FOR C(k)

Arcs Operating	Arcs Failed	Lower Bound[1] 1 Exact Term	Lower Bound[2] 2 Exact Terms	Upper Bound[3]
22	6	112861	192737	$\boxed{349618}$
23	5	23645	42484	64404
24	4	3754	7067	19506
25	3	423	$\boxed{827}$	3105
26	2	$\boxed{30}$[4]	$\boxed{30}$	

Notes: 1: Bounds obtained by projection using the value C(26) as known.

2: Bounds obtained by projection using the values C(26) and C(25) as known.

3: Bounds obtained using the number of trees as known.

4: Boxes indicate exact values obtained by enumeration of formula.

while the probability of a network having 24 operational links is $\binom{28}{24}(.05)^4(.95)^{24} = .037365$. We allow a total of 1000 samples allocated to 23 and 24 link nets in proportion to their probabilities of occurrence. Thus, we sample 24 link nets (1000) (.037365/ (.009439 + .037365) = 798.32 $\doteq$ 798 times and 23 link nets, (1000) (.009439/(.009439 + .037365) = 201.67 $\doteq$ 202 times. The results are illustrated in Table 2.3.3. along with the variance of the estimate for C(23) and C(24). In Table 2.3.4. upper and lower bounds for the probability of the network failing for p, between 0 and .1 in increments of .01 and from .1 to .9 in increments of .1 are given as well as an estimate of their value with the sample standard deviations. In general, we consider all the terms we don't know *a priori*. In our example, these are C(26), C(25), C(24) and C(23). To be specific, suppose we are willing to sample 1000 networks. For p = .05, the probability that a network with 26 links will occur is .24903; with 25 links, .11059; with 24 links, .03736; and with 23 links, .00944. We divide our 1000 samples between these 5 kinds of networks in proportion to their probability of occurring. This leads (approximately) to 608 samples of nets with 25 links and so on. However, there are only 378 networks with 26 links so it is more effective to enumerate C(26). This leaves 622 = 1000 − 378 samples to

TABLE 2.3.3

RESULTS OF SAMPLING STRATA

[illegible]	Number of Lots	NSAMP	μ	σ^2	σ	Est. No. of Disc. Lots
3	3276	440	.22954	.976	.987	751.97
4	20475	145	.45517	8.14	2.85	9319.60
5	98280	37	.64864	44.27	6.65	63748.33
4	20475	793	.48120	1.605	1.267	9852.57
5	[illegible]	202	.717821	8.64	2.9406	70547.44

TABLE 2.3.4

PROBABILITY FOR NETWORK BEING DISCONNECTED

AS A FUNCTION OF THE PROBABILITY OF LINK FAILURE

Link Failure Probability (p)	[illegible] Bound[1]	Lower Bound[2]	Upper Bound[1]	Estimate[3] By Sampling	Standard Deviation	Estimate[4] by Sampling	Standard Deviation
.01	.00267	.00301	.00311	.00297	.00005	.00303	.0
.02	.00957	.01186	.01321	.01178	.00032	.01221	.00003
.03	.01944	.02607	.03154	.02625	.00089	.02749	.00014
.04	.03159	.04465	.05888	.04613	.00174	.04865	.00036
.05	.04568	.05754	.09523	.07109	.00285	.07532	.00072
.06	.06177	.09361	.13985	.10075	.00417	.10704	.00121
.07	.08160	.12264	.19141	.13466	.00566	.14325	.00181
.08	.10123	.15436	.24824	.17230	.00729	.18330	.00250
.09	.12539	.18058	.30848	.21312	.00899	.22652	.00324
.1	.15298	.22511	.37035	.25654	.01392	.27220	.00517
.2	.57509	.64892	.84098	.70584	.01603	.72290	.00648
.3	.90281	.92194	.97578	.93931	.00530	.94493	.00214
.4	.98934	.99184	.99755	.99382	.00061	.99545	.00085
.5	.99954	.99963	.99987	.99971	.00002	.99974	.00001
.6	.99999	.99999	.99999	.99999	.0	.99995	.0
.7							
.8							
.9							

1. [illegible] were used for 3, 4, and 5 links failing.
2. [illegible] were used for 4 and 5 links failing.
3. [illegible] used for 3, 4, and 5 links failing.
4. [illegible] used for 4 and 5 links failing.

TABLE 2.3.5.

COMPARISON OF TWO DIFFERENT SIMULATION METHODS

Sample Size 1000

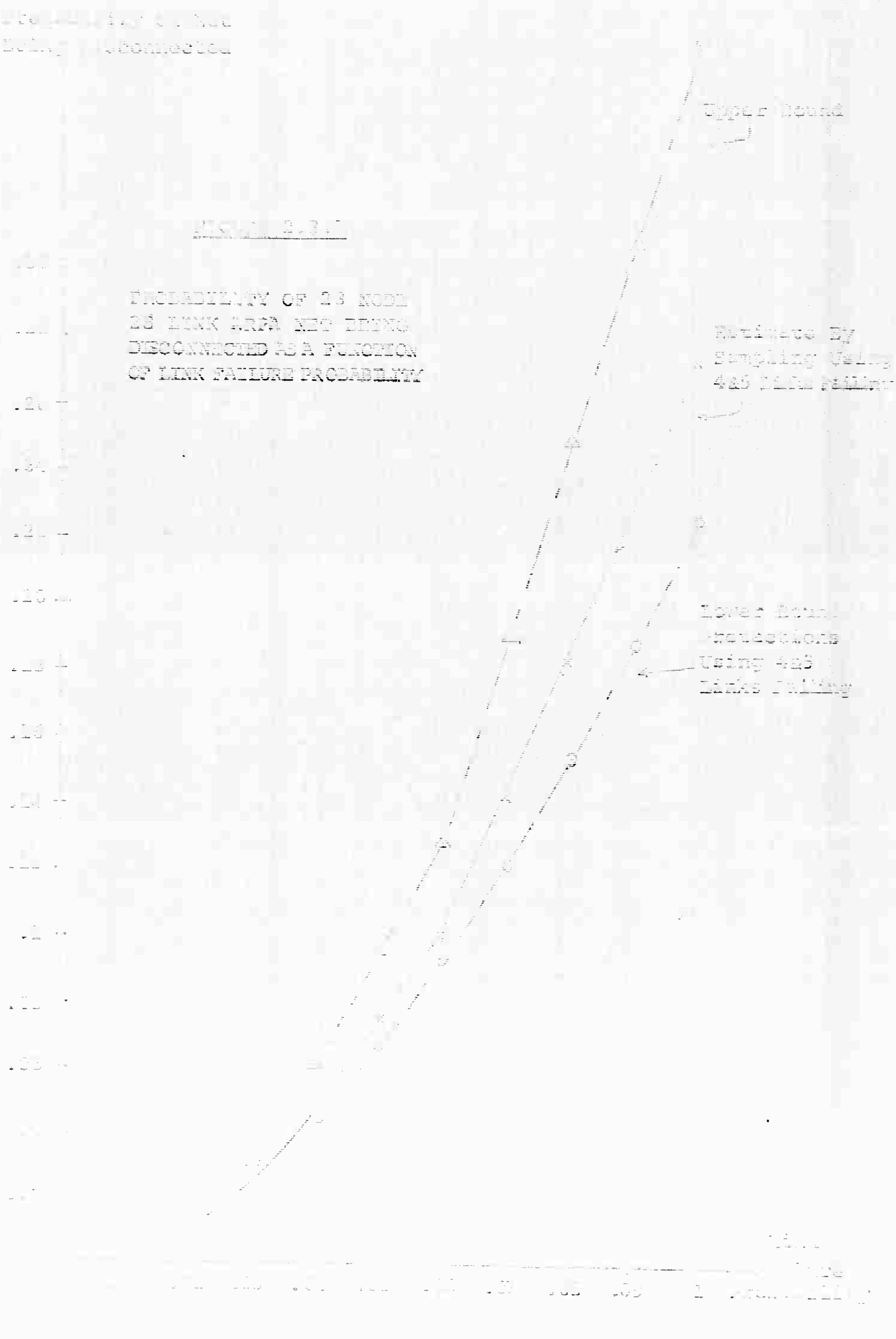
Link Failure Probability	Probability of Disconnected Net: Stratified Sampling	Probability of Disconnected Net: Straightforward Sampling	Standard Deviation: Stratified	Standard Deviation: Straightforward
.01	.00297	.004	5.15×10^{-5}	1.99×10^{-3}
.02	.01178	.012	3.28×10^{-4}	3.44×10^{-3}
.03	.02625	.027	8.96×10^{-4}	5.12×10^{-3}
.04	.04633	.042	1.74×10^{-3}	6.34×10^{-3}
.05	.07109	.070	2.85×10^{-3}	8.06×10^{-3}
.06	.10075	.097	4.17×10^{-3}	9.35×10^{-3}
.07	.13466	.135	5.66×10^{-3}	1.08×10^{-2}
.08	.17230	.176	7.29×10^{-3}	1.20×10^{-2}
.09	.21312	.224	8.99×10^{-3}	1.31×10^{-2}
.1	.25654	.276	1.07×10^{-2}	1.41×10^{-2}
.2	.70384	.743	1.60×10^{-2}	1.38×10^{-2}
.3	.93931	.954	5.30×10^{-3}	6.62×10^{-3}
.4	.99382	.995	6.32×10^{-4}	2.23×10^{-3}
.5	.99971	.999	2.89×10^{-5}	$1. \times 10^{-3}$
[illegible]	.99999	1.000	4.23×10^{-7}	0
.7				
.8				
.9				

allocate to networks with 23 through 25 operative links. Again, doing this proportionally we obtain 440 samples C(25), [illegible] for C(24) and 37 for C(23). The results are also displayed in Tables 2.3.3 and 2.3.4. In Table 2.3.5. the simulation using stratification with a sample of size 1000 is compared with conventional simulation with the same sample size done by assigning a random number to each link and considering the link failed or operative depending on whether the random number is less than or greater than the link failure probability.

Figure 2.3.1. illustrates the relationship between the upper and lower bounds given in the tables and the estimate of connectivity probability obtained by Simulation Method I.

2.4. AVERAGE FRACTION OF NON-COMMUNICATING NODE PAIRS

The determination of the expected number of node pairs which can communicate is done by sampling of randomly-generated networks themselves rather than of the coefficients C(K) in the formula for connectivity. This method, while less efficient, can be used in more general situations. In particular, it is easy to analyze the case of unequal failure probabilities. Basically, the method operates by generating a random number for each node and link. If the random number is less than the

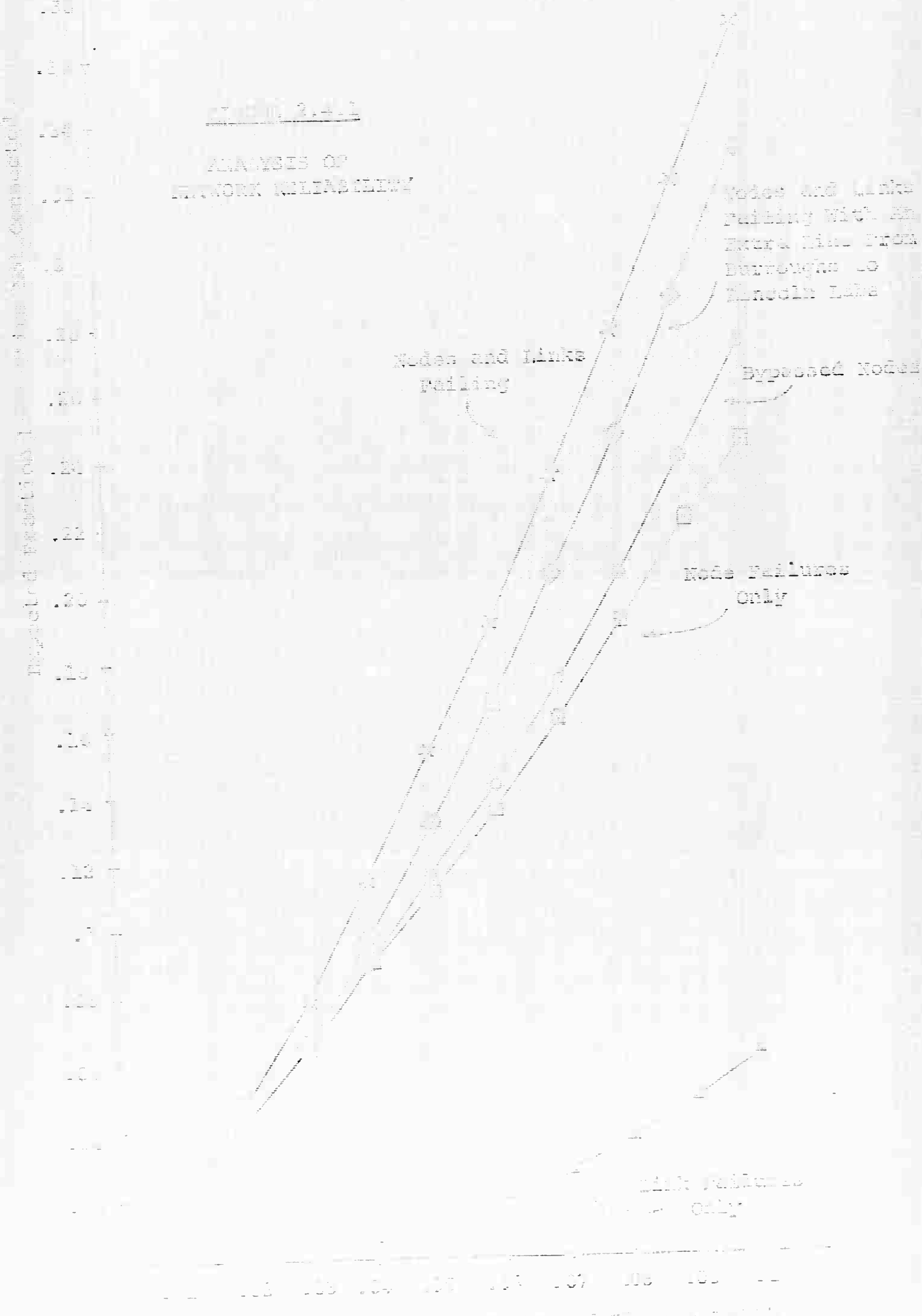


PROBABILITY OF 23 NODE 28 LINK ARPA NET BEING DISCONNECTED AS A FUNCTION OF LINK FAILURE PROBABILITY

failure probability, the corresponding node or link is failed, otherwise it is operative. We also use the simple device described in Section 1.4. which yields the expected number of pairs not communicating for a range of link and node failure probabilities.

For the first case, the results of a simulation on the 23 node, 28 link net in Figure 2.1.1, using equal probabilities for node and link failures and a sample of size 1000, are shown in Figure 2.4.1. Also shown in Figure 2.4.1. are the results of several experiments illustrating the flexibility of the simulation method. Three simulations were used to test ideas for improving the reliability of the network. The first idea was to remove the link connecting UTAH to NCAR and replacing it with a link from NCAR to WASH. The motivation for this change was to create an additional node disjoint path between the East and West Coasts. However, the difference as measured by the simulation was negligible and the two curves could not be shown separately in Figure 2.4.1. The next idea was to break up the long serial chain from BBN to HARVARD to BURROUGHS to ETAC to MITRE to CM to CASE. This chain is very vulnerable since any two node or link failures in the chain disconnects the network. To relieve this situation, a link was added [illegible]

ANALYSIS OF
NETWORK RELIABILITY



LINCOLN to BURROUGHS. This did cause enhancement [illegible] and is depicted in the figure. The third idea was to [illegible] hardware at the IMPs so that if an IMP failed, traffic could be routed around it in one direction connecting two of the incident links. Any remaining links are effectively blocked. In Figure 2.4.2. the directions of the bypasses chosen are depicted. This reduced the expected fraction of node pairs not communicating almost down to the level of the case of no-node failures. However, for low levels of unreliability, (p [illegible]) the improvement is not significant.

Two final simulations show what happens when only links fail and when only nodes fail. These simulations are interesting because they give some indication of the extent that any modification of the network design can lower the expected fraction of pairs communicating (EFPC). If a node fails, at least 21 pairs cannot communicate (all nodes paired with the failed node, independent of the network structure. Thus, good network design cannot improve the figure beyond the effect directly due to node failures. That this is rather small can be seen by comparing the EFPC for node failures only with the EFPC for link failures only.

The situation where link and node failure probabilities are not necessarily equal was also investigated.

23 Node Net

With Bypasses

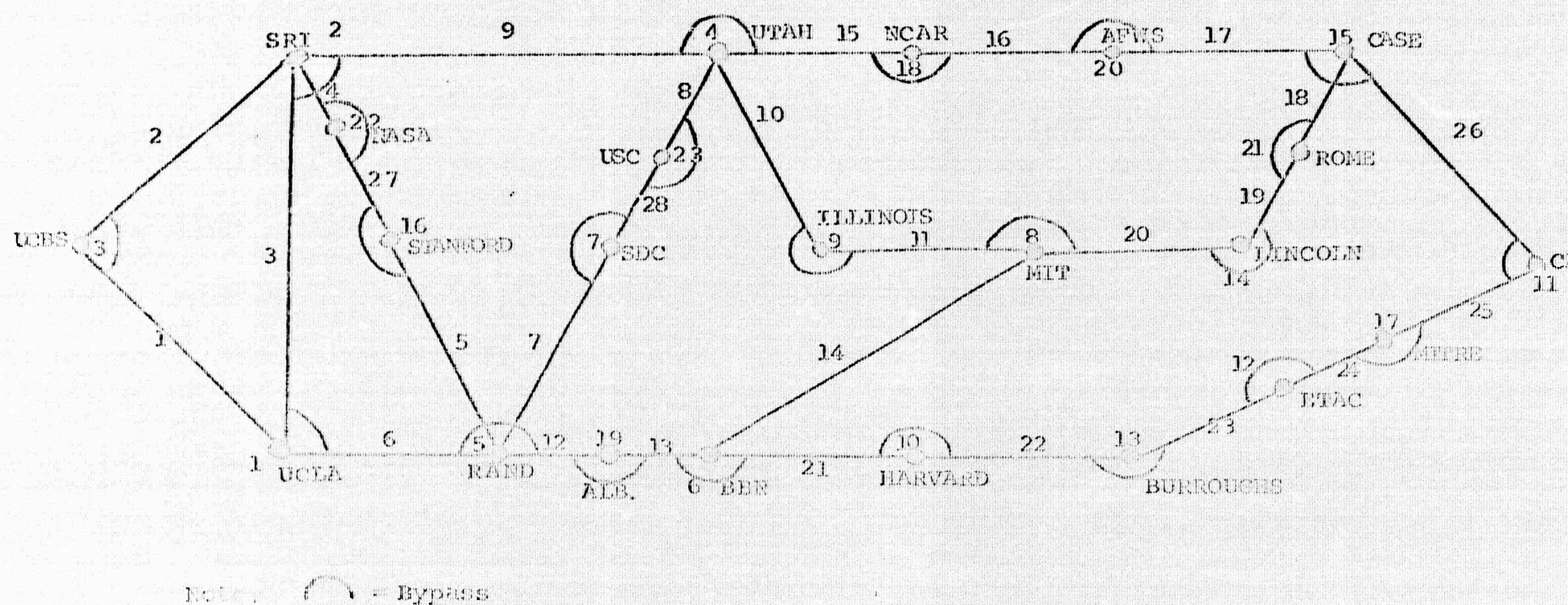


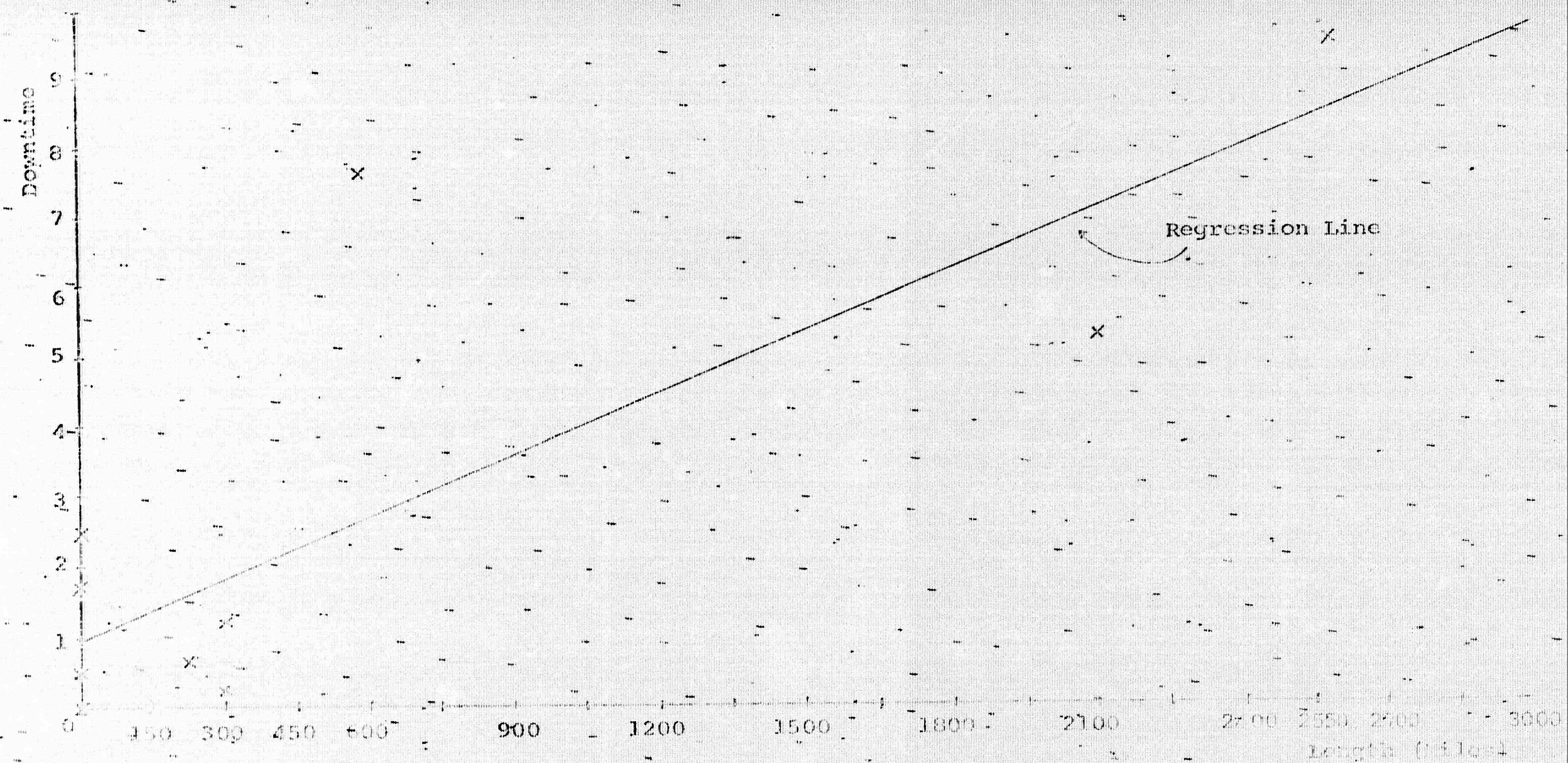
FIGURE 2.4.2

Preliminary data was available on the downtime for a subset of the communication lines on the ARPA network. The hypothesis was made that the reliability of a link in the ARPA net was a linear plus a constant term function of the link's length. Linear regression was used to fit a function to the available data. Table 2.4.1. shows the data used. The short links were all taken to be length 0 for the regression. The linear function chosen was .00293 X + .904 which gives the percent downtime as a function of the direct distance X in miles between nodes. The fit of the regression is displayed in Figure 2.4.3. A failure probability of .03 was assigned to the nodes. The failure probabilities for the links obtained using the regression equation are displayed in Table 2.4.2. The average failure probability over all the nodes and links was .0241. In Figure 2.4.4. the results of the simulation are displayed.

Additionally, some points plotted were obtained from Figure 2.4.1. by averaging all the link and node probabilities to obtain a common failure probability. From the close similarity, we conclude that for design purposes the assumption that links fail with equal probability is a good first approximation.

TABLE 2.4.1

Line	Nominal Inst. Date	Approx. Length	Number of Failures	% Downtime
RAND-BBN(12)	6/1/70	2600	22	9.44
UTAH-MIT(11)	6/15/70	2100	58	5.31
SRI-UTAH(10)	6/1/70	600	7	1.96
SDC-UTAH(9)	6/1/70	580	23	7.59
STAN-RAND(8)	7/15/70	300	1	.18
UCLA-SRI(7)	6/1/70	300	10	1.17
SRI-USCB(6)	6/1/70	225	2	.59
USCB-UCLA(5)	6/1/70	150	0	0
UCLA-RAND(4)	6/1/70	Short	2	1.76
STAN-SRI(3)	7/15/70	Short	1	.46
RAND-SDC(2)	6/1/70	Short	1	.036
BBN-MIT(1)	6/1/70	Short	11	2.51



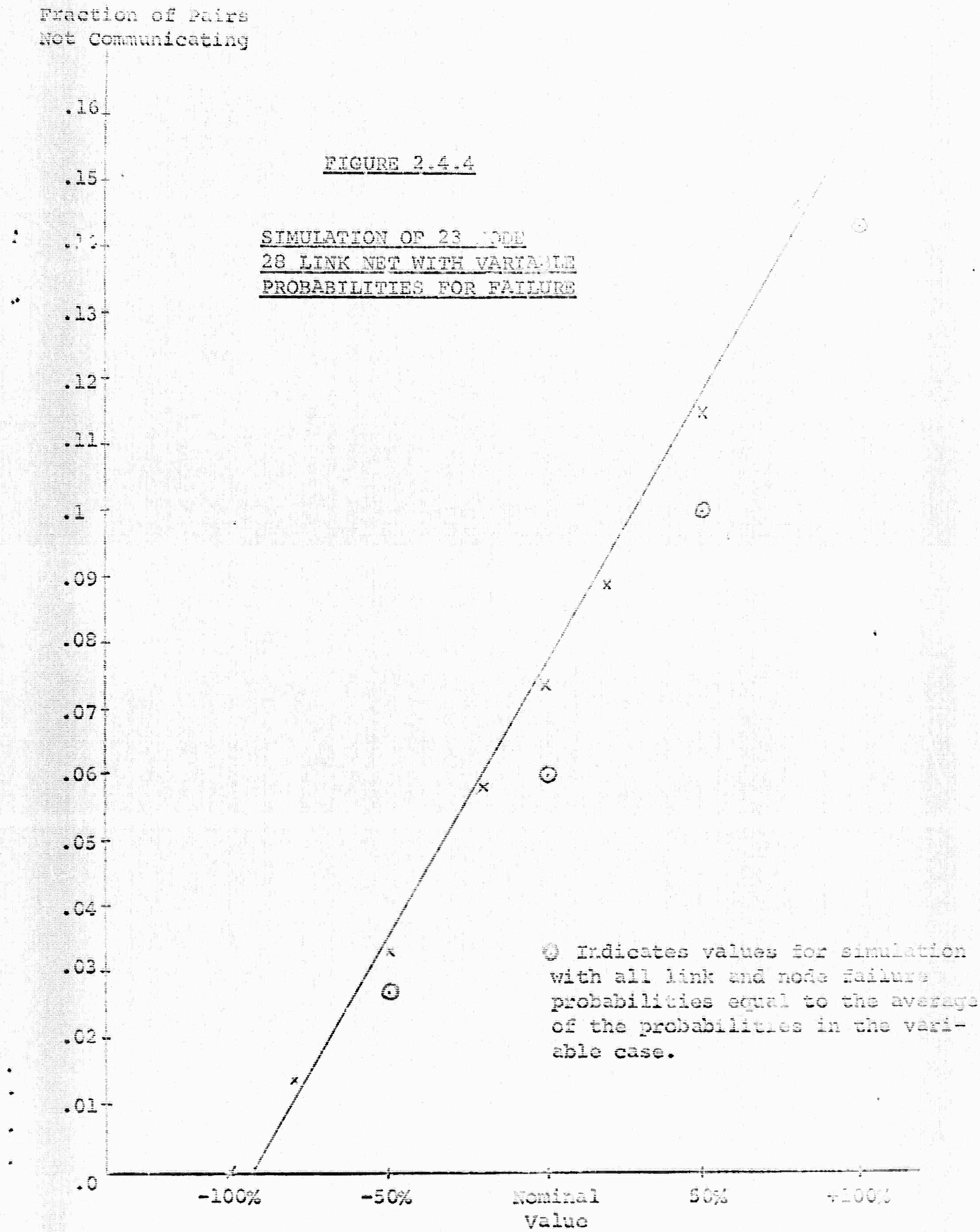
PERCENT DOWNTIME VS. DISTANCE FOR [illegible] LINES

FIGURE 2.4.3

TABLE 2.4.2

LINK FAILURE PROBABILITIES BASED ON REGRESSION EQUATION

Link No.	From	To	Failure Probability
1	UCLA	UCSB	.0105
2	SRI	UCSB	.0162
3	UCLA	SRI	.0176
4	SRI	NASA	.0090
5	RAND	STANFORD	.0180
6	UCLA	RAND	.0096
7	RAND	SDC	.0180
8	UTAH	USC	.0257
9	SRI	UTAH	.0255
10	UTAH	ILLINOIS	.0439
11	MIT	ILLINOIS	.0345
12	RAND	ALB	.0265
13	BBN	ALB	.0633
14	BBN	MIT	.0090
15	UTAH	NCAR	.0189
16	NCAR	AFWS	.0228
17	CASE	AFWS	.0300
18	CASE	ROME	.0188
19	LINCOLN	ROME	.0156
20	MIT	LINCOLN	.0090
21	BBN	HARVARD	.0090
22	HARVARD	BURROUGHS	.0166
23	ETAC	BURROUGHS	.0126
24	ETAC	MITRE	.0096
25	CM	MITRE	.0144
26	CM	CASE	.0123
27	STANFORD	NASA	.0090
28	SDC	USC	.0181


Fraction of Pairs
Not Communicating
.16
.15
.14
.13
.12
.11
.1
.09
.08
.07
.06
.05
.04
.03
.02
.01
.0
FIGURE 2.4.4
SIMULATION OF 23 NODE
28 LINK NET WITH VARIABLE
PROBABILITIES FOR FAILURE
Indicates values for simulation with all link and node failure probabilities equal to the average of the probabilities in the variable case.
-100%
-50%
Nominal
Value
50%
+100%

It is also possible to use the Criterion C1 for networks with failing nodes and links. In Figure 2.4.5 the probability that the network is disconnected is compared for three situations. In the first case, nodes and links have the same failure probability and the network is considered disconnected if any node pair cannot communicate. In particular, if any node has failed, the network is counted as disconnected. In the second case, we again have equal node and link failure probabilities, but the network is considered disconnected only if a pair of _operable_ nodes cannot communicate. Finally, for comparison purposes we give the probability of disconnection for the case of no node failures.

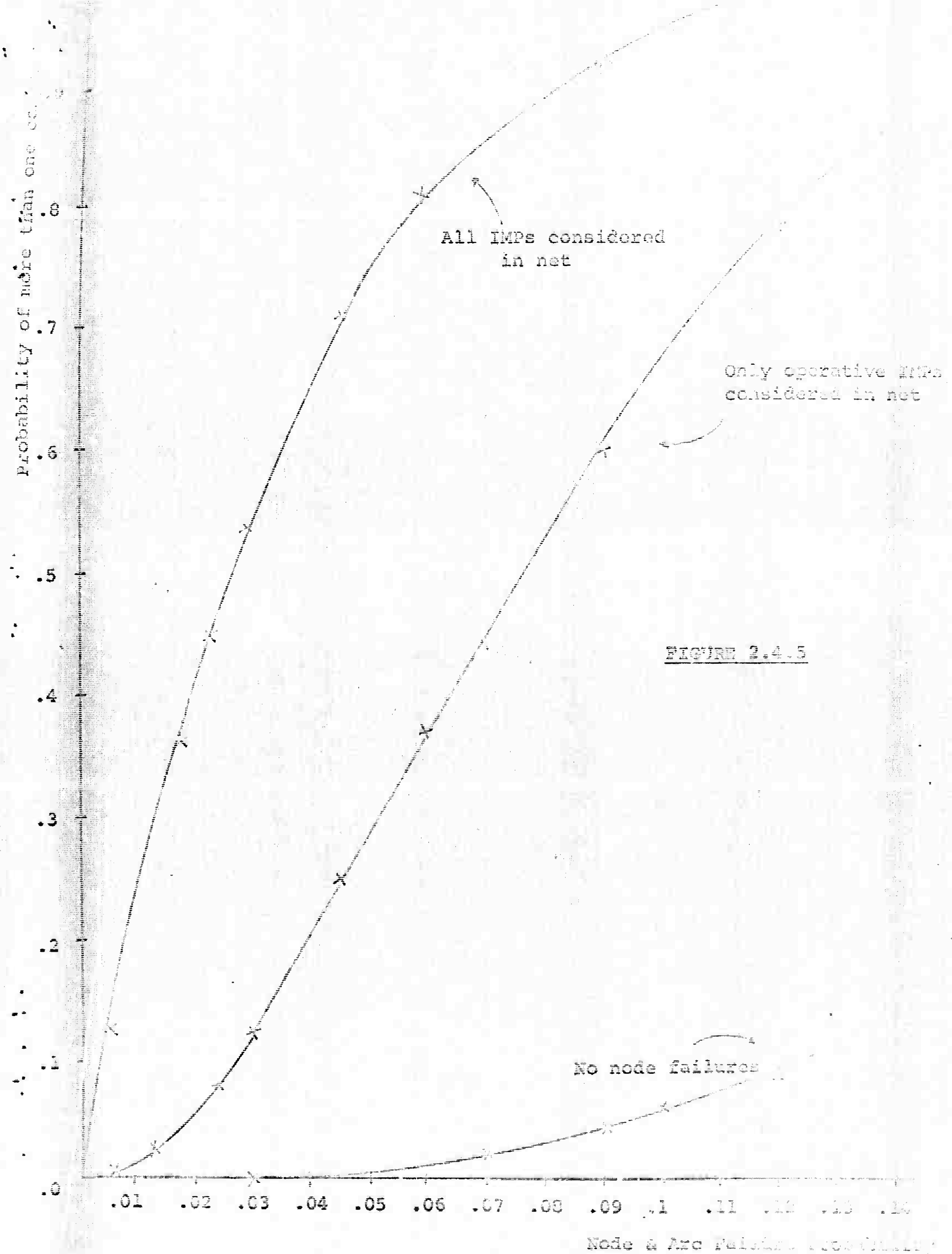

All IMPs considered
in net
No node failures
.8
.7
.6
.5
.4
.3
.2
.1
.0
.01
.02
.03
.04
.05
.06
.07
.09
.1
.11

FIGURE 2.4.5

2.5 FUTURE DEVELOPMENTS

The criterion most frequently used by NAC in comparing computer network design reliability is the percentage of node pairs communicating. This can also be interpreted as the average fraction of the other nodes with which an average node can communicate. This is a straightforward generalization of our earlier techniques.

In the earlier work on designing the ARPA network, network cost was minimized subject to delay time constraints and the constraint that the network be "two-connected." (That is, the constraint that every node pair be connected by at least two node disjoint paths.) This last requirement was to guarantee a minimum level of reliability for the network. Now, as a result of the research reported in the previous sections, more sophisticated analysis of the reliability of the ARPA network are possible. Future research will follow two directions. The first line of approach is to investigate the reliability of networks designed with two node disjoint paths for every node pair as a function of size. The reliability of ARPA-like nets designed to connect respectively 20, 40, 60, 80, 100 and 200 IMPS will be computed using special decomposition methods especially developed for analyzing large nets. The second approach

will be to design networks including direct, explicit constraints on the reliability rather than on connectivity. This approach represents a major advance in network reliability and survivability.

3. SPECIALIZED NETWORK STUDIES

3.1. INTRODUCTION

This chapter describes some specialized studies of cost and throughput trade-offs for the ARPA Network and for larger store-and-forward systems. Section 3.2 summarizes the results of a number of topological design studies for the ARPA Network. Section 3.3 provides a preliminary analysis of the costs of providing 80 KB/Sec. input capacity to a percentage of the nodes in 20, 40, 60, 80, and 100 node networks. Section 3.4 documents an initial investigation of the costs of centralizing store-and-forward capabilities in telephone company switching offices while Section 3.5 provides two designs for a 200 node store-and-forward network.

3.2. TOPOLOGICAL OPTIMIZATION OF ARPA NETWORK STRUCTURE

During the course of the reporting period, a number of specialized optimizations have been performed to introduce new nodes into the ARPA Network. This section summarized these runs.

Figures 3.2.1.(a) - 3.2.1.(j) show a sequence of ARPA Network designs as additional nodes are added to the network. Table 3.2.1 gives the coordinates of the nodes while Table 3.2.2 shows the relative merits of each of these systems. Figure 3.2.2

indicates the relationship between the cost per node and the number of nodes in the evolving ARPA Network as given in Table 3.2.[illegible].

Figures 3.2.3(a) - (c) show the networks derived on the basis of 21 node network optimization. These results are summarized in Table 3.2.3. Figure 3.2.4 shows such an optimization for a 24 node ARPA Network while Figure 3.2.5 shows results for a 26 node optimization. Tables 3.2.4 and 3.2.5 summarize the various performance characteristics for the 24 and 26 node network optimizations, respectively.

Finally, Figure 3.2.6 shows the result of a different optimization problem--the addition of a group of 7 nodes representing the University of California System. (UCLA and UCSB are already in the Network.) The cost performance characteristics of the network shown in Figure 3.2.6 are given in Table 3.2.6.

TABLE 3.2.1

NODE COORDINATES

Node Number	Node Name	Node Location Latitude	Node Location Longitude
1	UCLA	34 04	118 31
2	SRI	37 22	122 10
3	UCSB	34 30	119 45
4	UTAH	40 40	111 50
5	RAND	34 00	118 35
6	BBN	42 30	71 20
7	SDC	34 01	118 33
8	MAC	42 30	71 12
9	ILLINOIS	40 05	88 30
10	HARVARD	42 30	71 15
11	CARNEGIE-MELLON	40 30	79 50
12	ETAC (WASHINGTON)	38 50	77 00
13	PAOLI	38 55	77 10
14	LINCOLN LABS	42 35	71 20
15	CASE	41 30	81 45
16	STANFORD	37 18	122 10
17	MITRE	39 00	77 00
18	NCAR DENVER	39 30	105 00
19	ALBURQUERQE	35 05	106 40
20	AFWS OMAHA	41 00	96 00
21	ROME, N.Y.	43 15	75 25
22	NASA	37 17	122 02
23	USC	34 00	118 21
24	TINKER	35 27	97 32
25	McCLELLAN	38 35	121 30
26	NBS	39 08	77 10

TABLE 3.2.2

NETWORK CHARACTERISTICS

Fig.3.2.1( )	Number of Nodes	Yearly Line Cost (K$)	Throughput KBS/NODE Uniform Traffic	Line Cost/Node (K$)	Line Cost/ Megabit (¢)
(a)	10	524	19	52.4	8.77
(b)	14	605	10.5	43.2	13.00
(c)	15	659	10.7	43.9	13.02
(d)	18	792	12.2	44.0	11.46
(e)	21	825	10.6	39.3	11.71
(f)	21	825	10.4	39.3	11.95
(g)	23	847	9.9	36.8	11.81
(h)	23	849	10.2	36.9	11.51
(i)	24	860	9.5	35.8	11.90
(j)	26	883	8.6	34.0	12.48

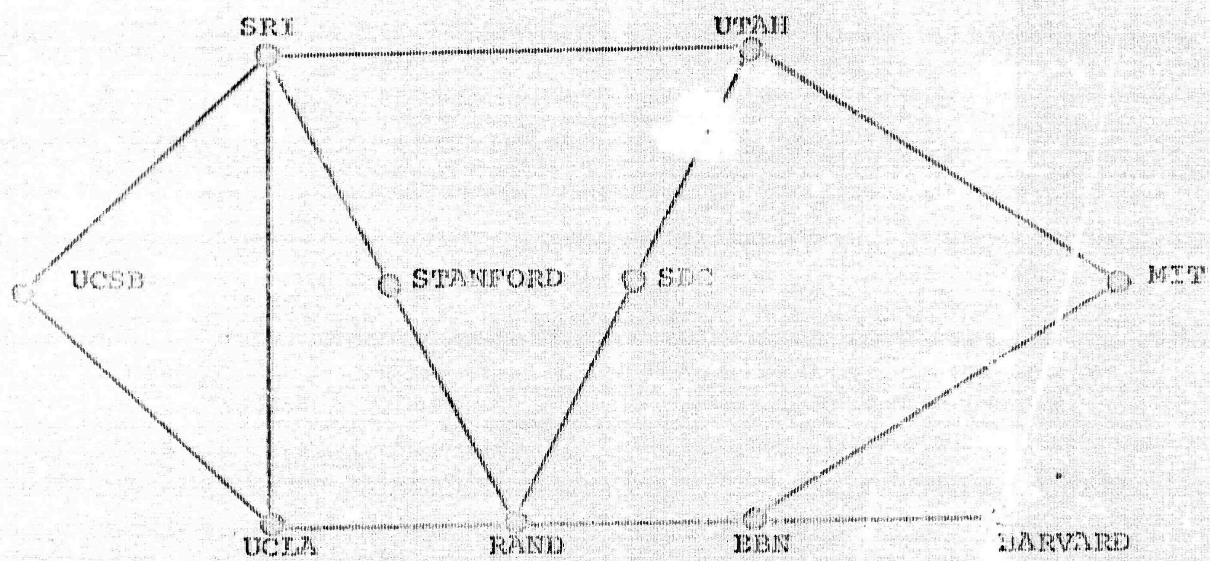


FIGURE 3.2.1 (a)

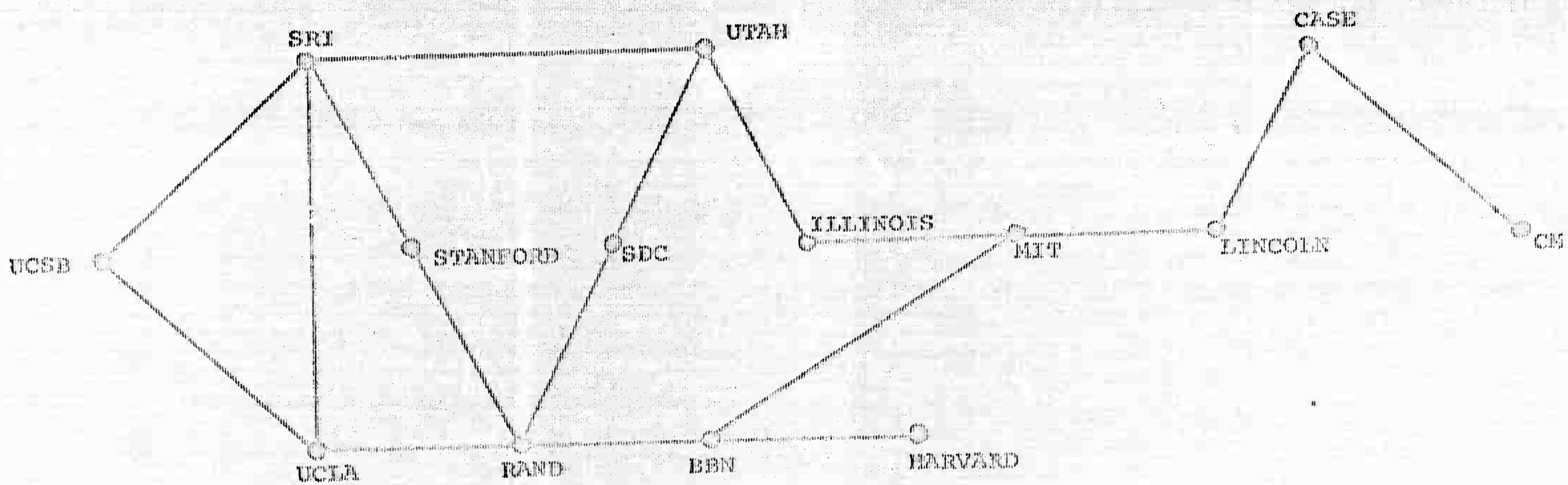

SRI
UTAH
CASE
UCSB
STANFORD
SDC
ILLINOIS
MIT
LINCOLN
CM
UCLA
RAND
BBN
HARVARD

FIGURE 3.2.1 (b)

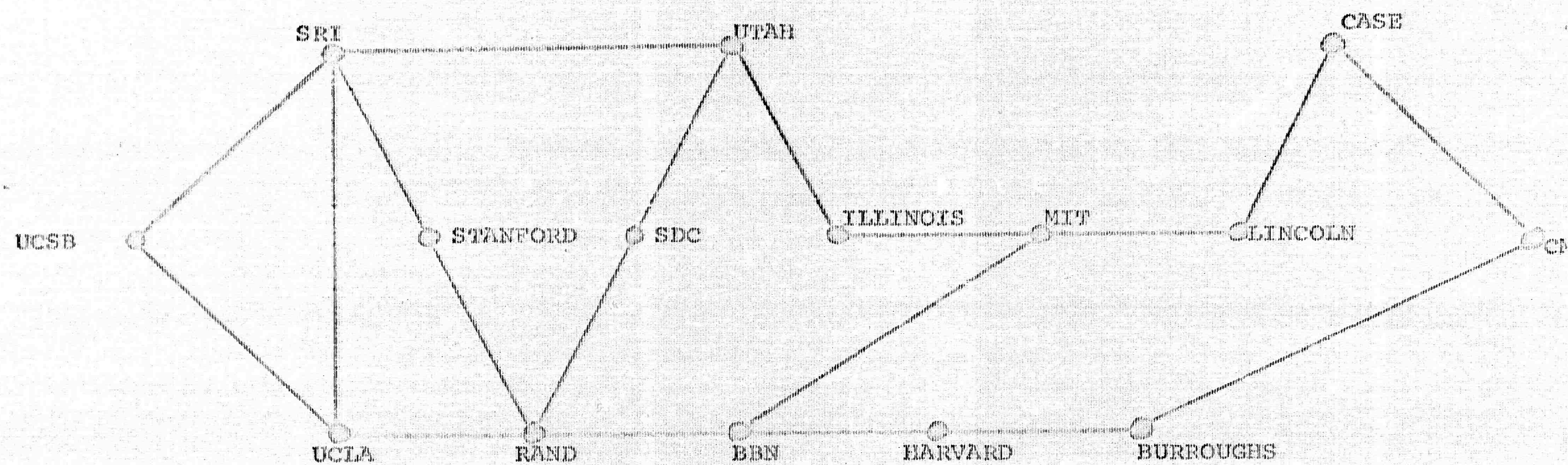

SRI
UTAH
CASE
UCSB
STANFORD
SDC
ILLINOIS
MIT
LINCOLN
CM
UCLA
RAND
BBN
HARVARD
BURROUGHS

FIGURE 3.2.1 (c)

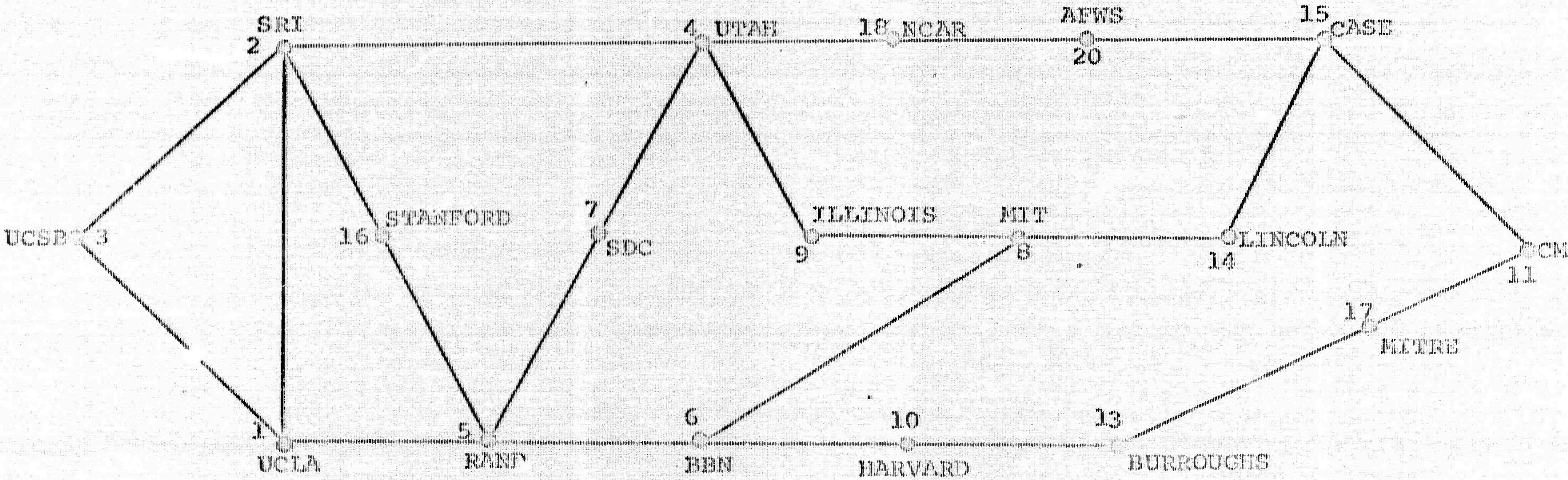

SRI
2
4 UTAH
18 NCAR
AFWS
20
15 CASE
UCSB 3
16 STANFORD
7
SDC
ILLINOIS
9
MIT
8
LINCOLN
14
CM
11
17
MITRE
1
UCLA
5
RAND
6
BBN
10
HARVARD
13
BURROUGHS

FIGURE 3.2.1 (d)

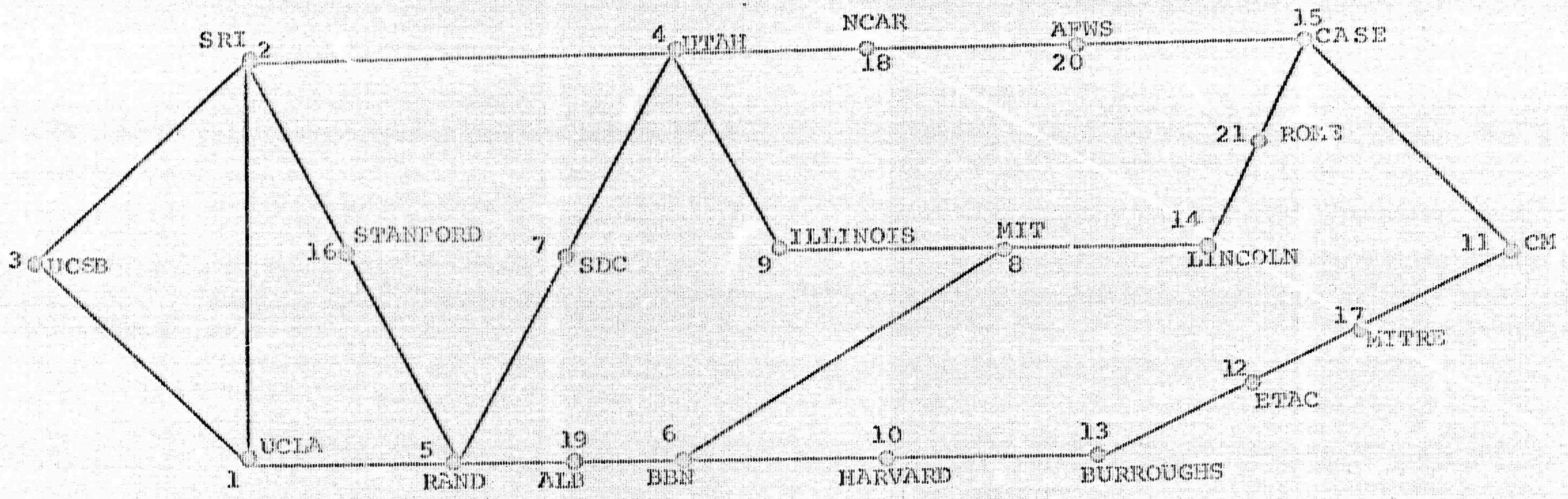

SRI 2
3 UCSB
1 UCLA
16 STANFORD
5 RAND
7 SDC
4 UTAH
19 ALB
6 BBN
9 ILLINOIS
NCAR 18
AFWS 20
MIT 8
10 HARVARD
14 LINCOLN
21 ROME
15 CASE
13 BURROUGHS
12 ETAC
17 MITRE
11 CM

FIGURE 3.2.1 (e)

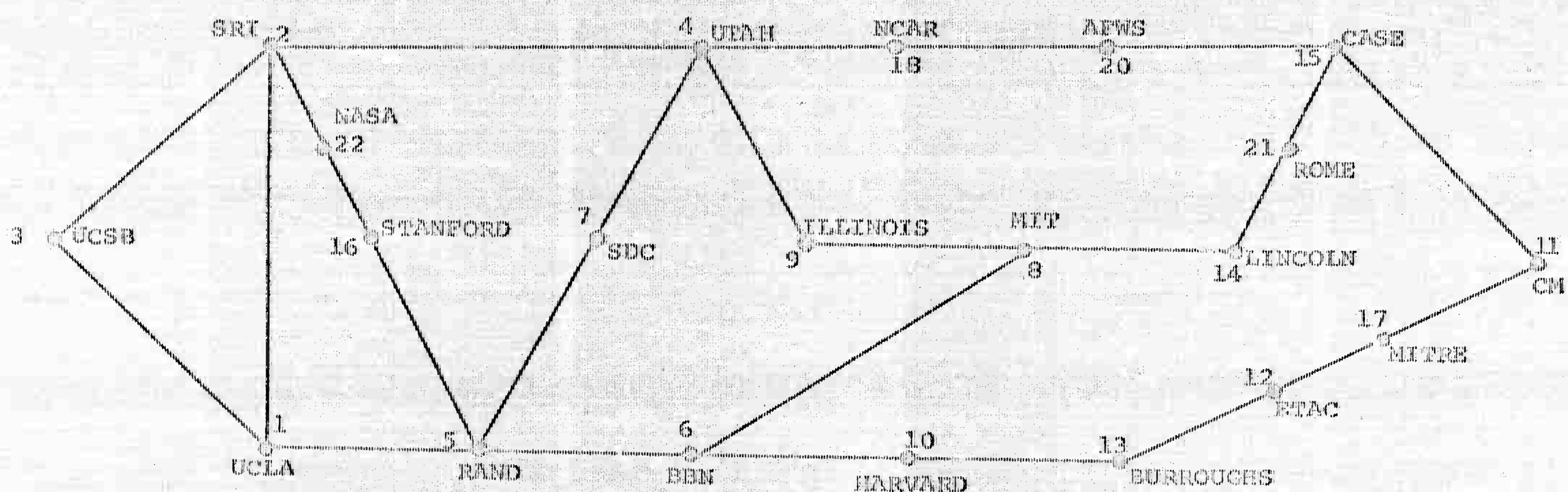
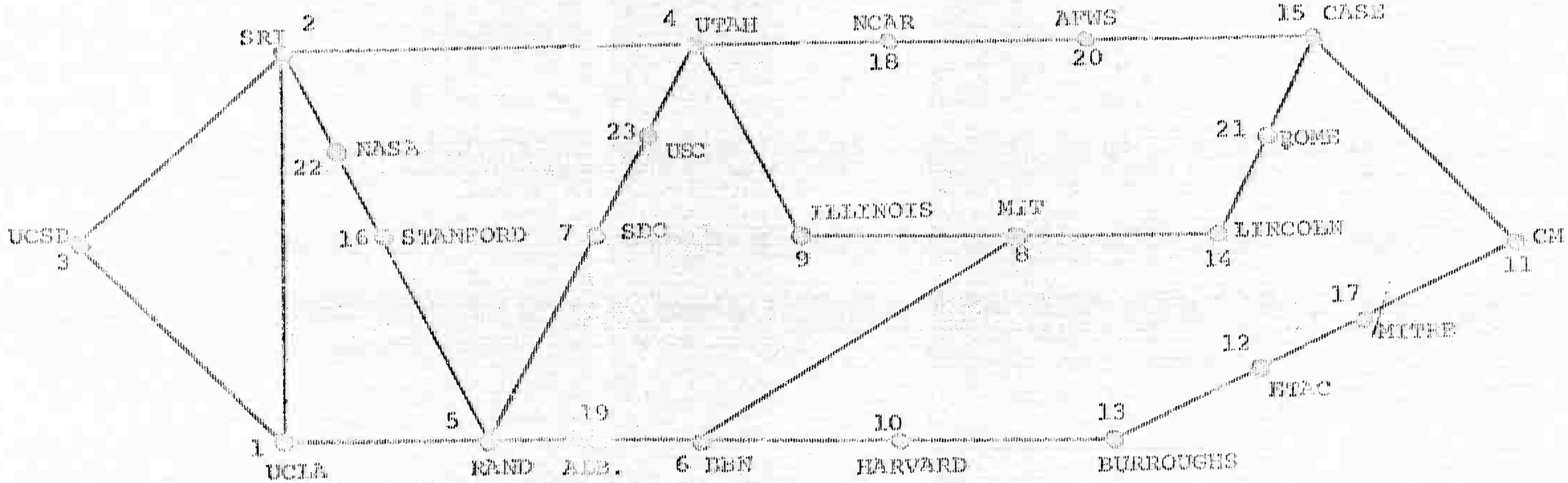

SRI 2
4 UTAH
NCAR
18
AFWS
20
15 CASE
NASA
22
21
ROME
3 UCSB
16 STANFORD
7
SDC
ILLINOIS
9
MIT
8
14 LINCOLN
11
CM
17
MITRE
12
ETAC
1
UCLA
5
RAND
6
BBN
10
HARVARD
13
BURROUGHS

FIGURE 3.2.1 (F)



SRI 2
4 UTAH
NCAR 18
AFWS 20
15 CASE
21 ROME
23 USC
22 NASA
UCSD 3
16 STANFORD
7 SDC
ILLINOIS 9
MIT 8
LINCOLN 14
CM 11
17 MITRE
12 ETAC
1 UCLA
5 RAND
19 ALB.
6 BBN
10 HARVARD
13 BURROUGHS

FIGURE 3.2.1 (g)

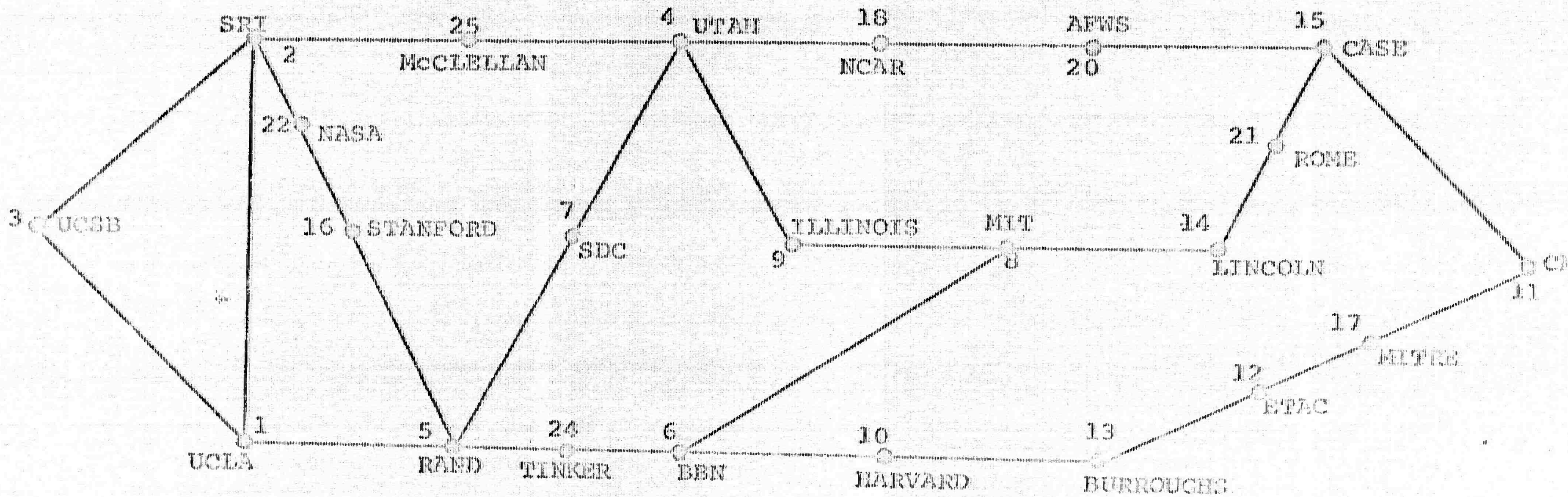

SRI
2
25
McCLELLAN
4
UTAH
18
NCAR
AFWS
20
15
CASE
22
NASA
21
ROME
3
UCSB
16
STANFORD
7
SDC
9
ILLINOIS
MIT
8
14
LINCOLN
11
17
MITRE
12
ETAC
1
UCLA
5
RAND
24
TINKER
6
BBN
10
HARVARD
13
BURROUGHS

FIGURE 3.2.1 (b)

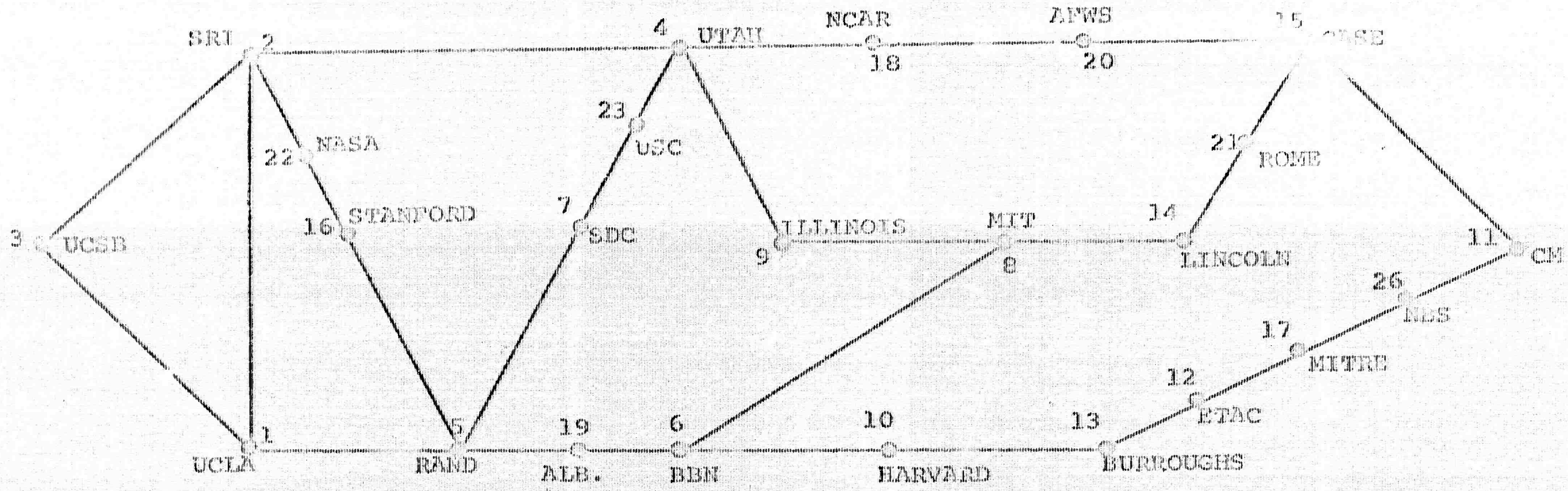
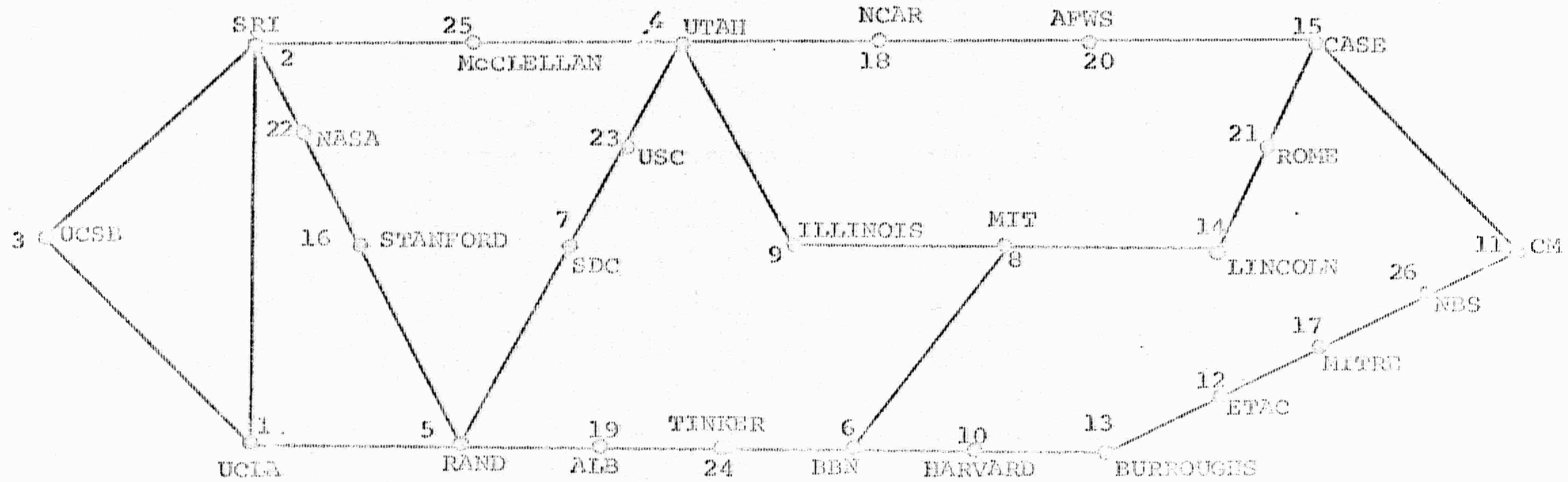
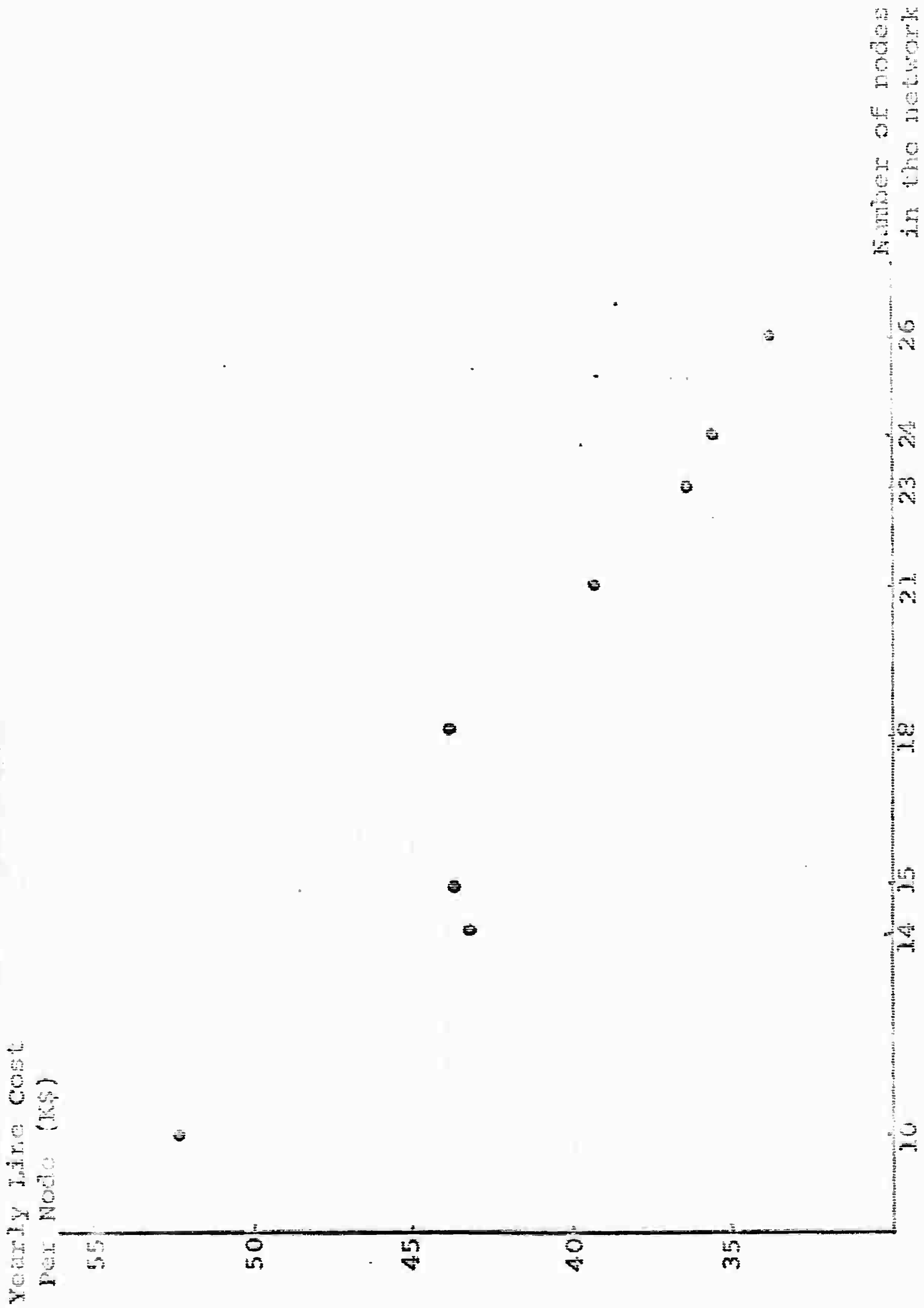

SRI
2
UCSB
3
UCLA
1
22
NASA
16
STANFORD
RAND
5
23
USC
7
SDC
4
UTAH
19
ALB.
6
BBN
NCAR
18
APWS
20
9
ILLINOIS
MIT
8
10
HARVARD
15
21
ROME
14
LINCOLN
13
BURROUGHS
12
ETAC
17
MITRE
26
NBS
11
CM

FIGURE 3.2.1 (i)



SRI
2
25
McCLELLAN
4
UTAH
NCAR
18
AFWS
20
15
CASE
22
NASA
23
USC
21
ROME
3
UCSB
16
STANFORD
7
SDC
9
ILLINOIS
MIT
8
14
LINCOLN
11
CM
26
NBS
17
MITRE
12
ETAC
1
UCLA
5
RAND
19
ALB
TINKER
24
6
BBN
10
HARVARD
13
BURROUGHS

FIGURE 3.2.1 (5)



RELATIONSHIP BETWEEN COST AND SIZE
FOR THE EVOLVING ARPA NETWORK

FIGURE 3.2.2

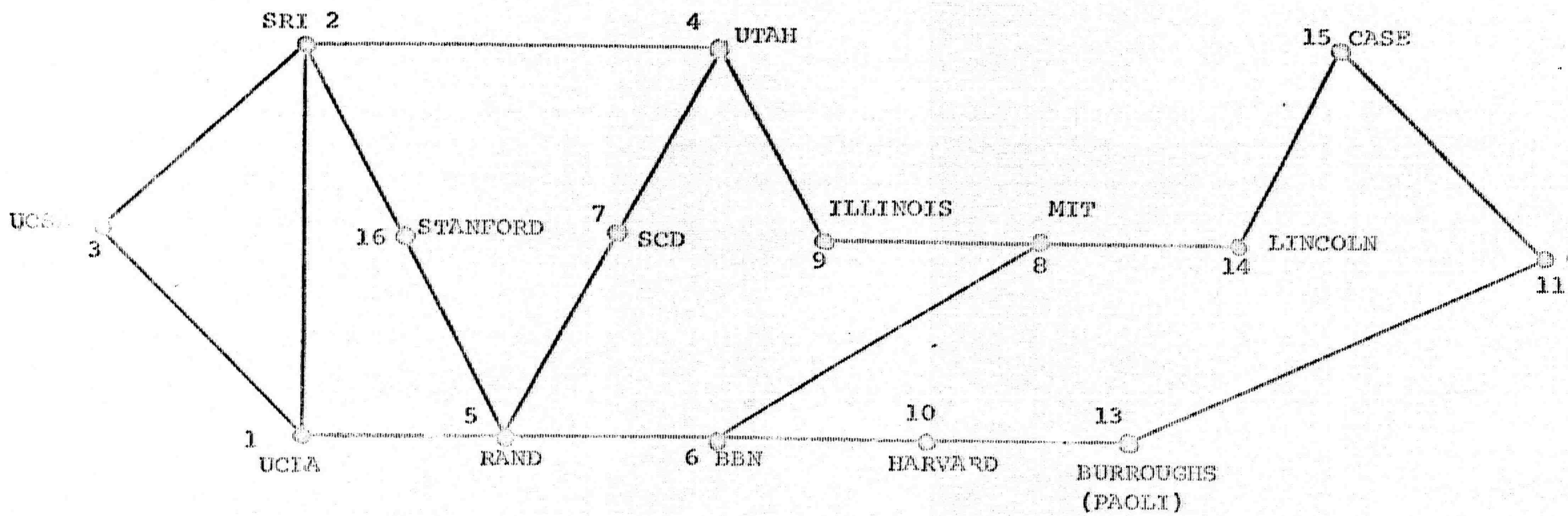
TABLE 3.2.3

21 NODE OPTIMIZATION

Assumptions:
1. 85% of messages short.
2. Uniform traffic between nodes.
3. Cheap imps cost \$17.6K + \$500K/N
4. Expensive imps cost \$33.3K + \$500K/N

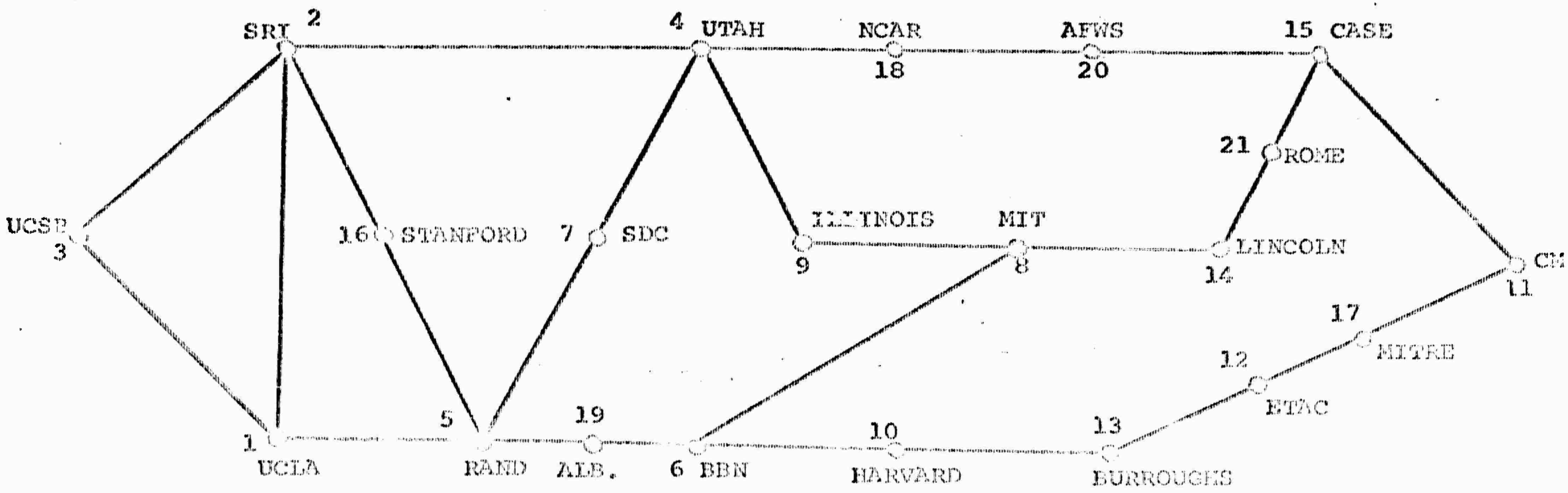
Note: All Lines have 50 KB/Sec Capacity at a Monthly Cost of \$850 + \$5.00/mile

	KBITS/SEC/NODE	YEARLY LINE COST	YEARLY TOTAL COST	COST PER NODE	ADDED LINE
1.	11.0	\$ 825,173	\$ 1,430,173	\$ 68.1K	Network shown in Figure 3.2.2(b). This network was derived from the Janu[illegible] 1971 configur[illegible] shown in Figure [illegible].2.2(a).
2.	12.0	\$ 858,870	\$ 1,463,700	\$ 69.7K	(13,14)
3.	14.0	\$ 1,018,200	\$ 1,623,200	\$ 77.3K	(3,10)
4.	16.0	\$ 1,147,700	\$ 1,742,700	\$ 83.5K	(7,11)
5.	17.0	\$ 1,192,700	\$ 1,797,700	\$ 85.7K	(9,17)
6.	18.0	\$ 1,222,000	\$ 1,827,000	\$ 86.6K	([illegible])
7.	19.0	\$ 1,251,500	\$ 1,856,500	\$ 88.4K	(18,19)



15 Node January, 1971 ARPA CONFIGURATION

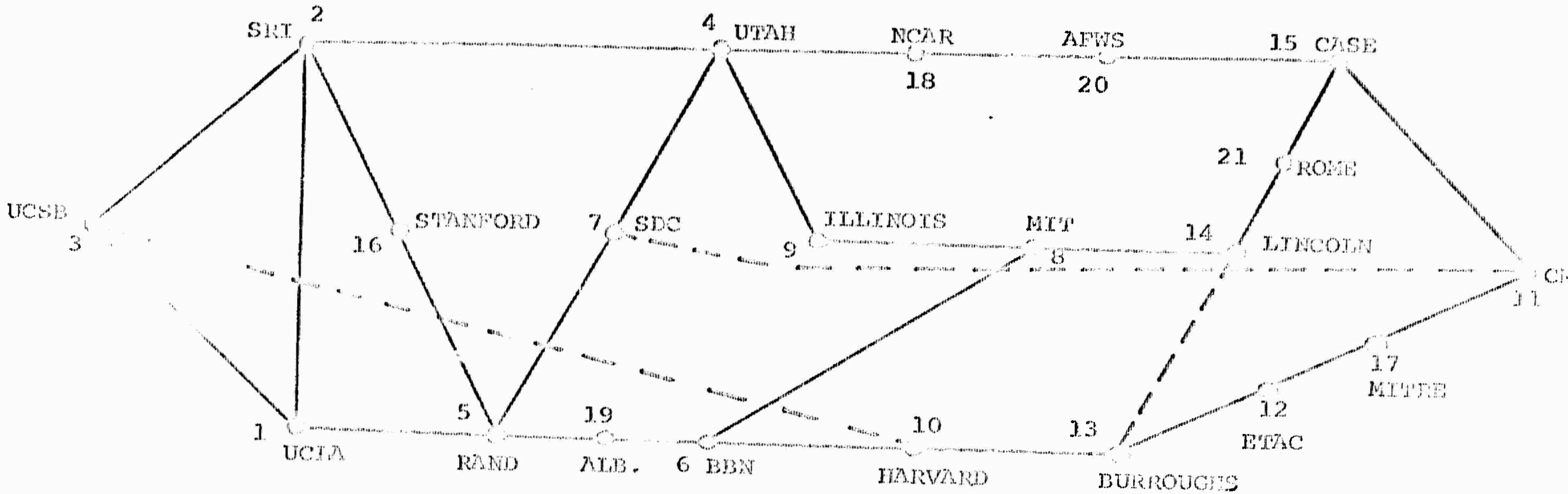
FIGURE 3.2.3(a)



Base Network for Optimization
Derived from Figure 3.2.?(a)

FIGURE 3.2.3 (b)

all lines 50 kilobits/sec



SUMMARY OF LINES ADDE BY OPTIMIZATION

FIGURE 3.2.3.(c)

TABLE [illegible]

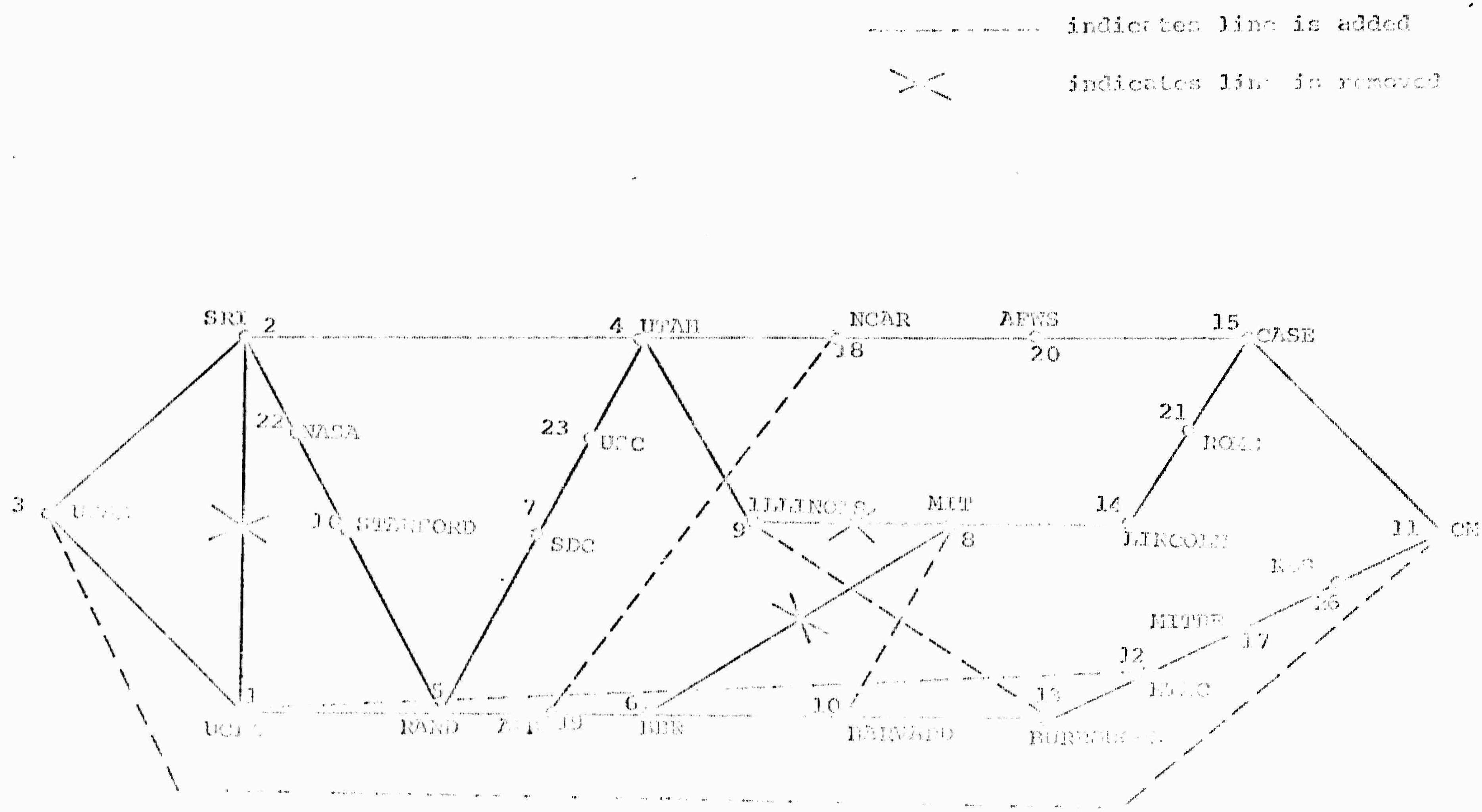
24-NODE OPTIMIZATION

Throughput (KBPS/NODE)	Yearly Line Cost ($K)	Line Cost Per Node	Line Cost per Megabit (Cents)	[illegible]	[illegible]
9.5	860	35.8	11.90	Network shown in Figure 3.[illegible](a)	
9.5	832	34.7	11.50		([illegible], 2)
12.0	968	40.0	10.06	(1, 12)	
14.5	1,101	46.0	10.00	(3, 11)	
17.0	1,112	46.4	8.60	(18, 19) (6, 10) (9, 13)	([illegible], 6) (5, 8)

TABLE 3.2.5

26-NODE OPTIMIZATION

Throughput (KBPS/NODE)	Yearly Line Cost (K$)	Line Cost Per Node	Line Cost per Megabit (Cents)	Network Description Added	Network Description Removed
8.6	883	34.0	12.48	Network shown in Fig. 3.2.1 (j)	
8.6	855	32.9	12.20		(1, 2)
10.5	991	38.0	11.60	(1, 12)	
13.0	1129	43.5	10.07	(2, 11)	
14.0	1144	44.0	10.00	(3, 23)	
16.0	1173	45.1	9.00	(1, 7) (3, 22) (9, 8) (24, 20)	(1, 2) (3, 1) (9, 6)


indicates line is added
indicates line is removed
SRI
2
4
UTAH
NCAR
18
AFWS
20
15
CASE
22
NASA
23
UTC
21
3
10
STANFORD
7
SDC
ILLINOIS
9
MIT
8
14
LINCOLN
11
MITRE
17
12
1
5
RAND
6
BBN
10
HARVARD
13

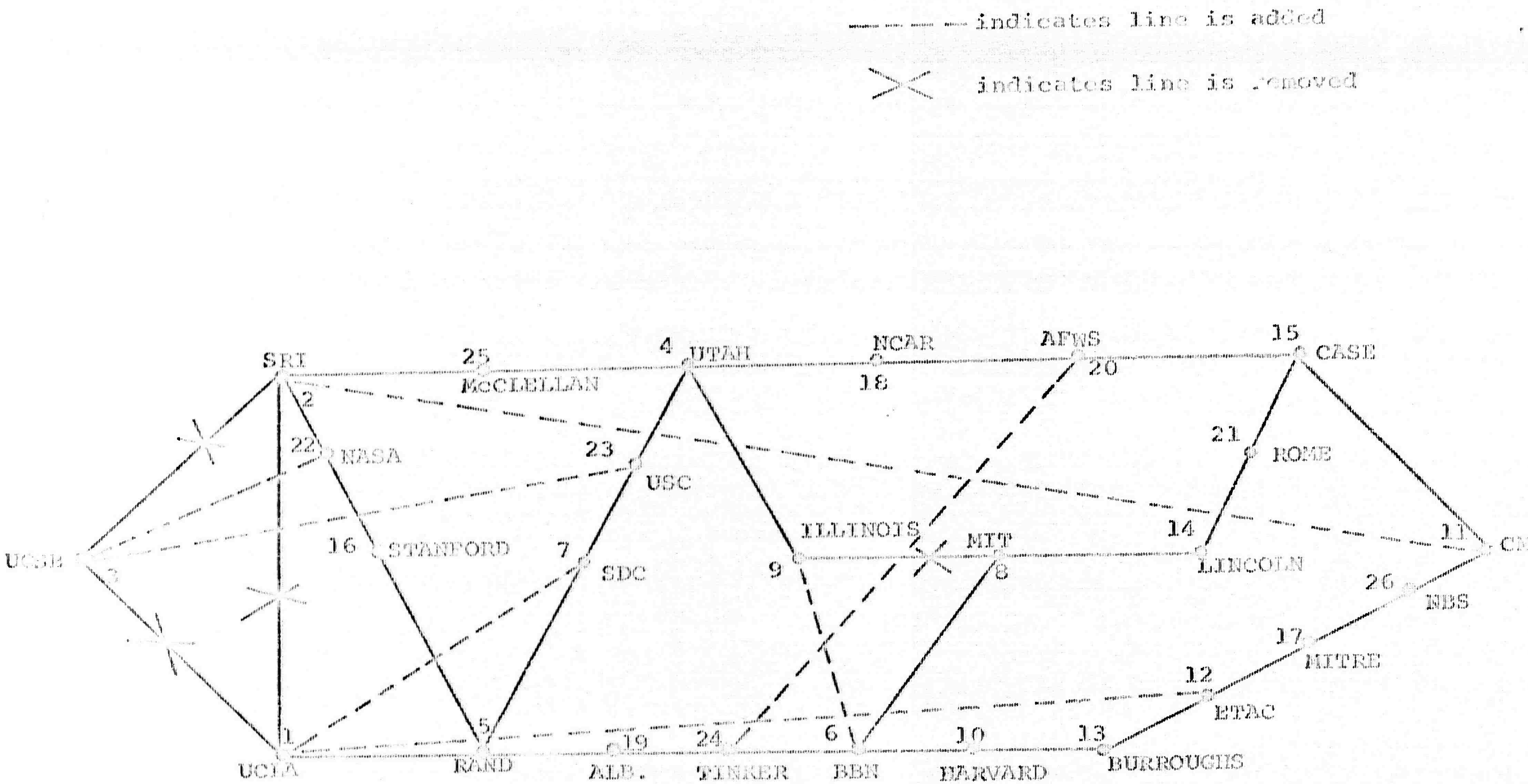

indicates line is added
indicates line is removed
UCSB
3
SRI
2
22
NASA
16
STANFORD
1
UCLA
25
McCLELLAN
4
UTAH
23
USC
7
SDC
5
RAND
19
ALB.
24
TINKER
NCAR
18
AFWS
20
ILLINOIS
9
6
BBN
MIT
8
10
HARVARD
13
BURROUGHS
15
CASE
21
ROME
14
LINCOLN
11
CMU
26
NBS
17
MITRE
12
ETAC

FIGURE 3.2.5

TABLE 3.2.6

COMPARISON OF ARPA NETS WITH AND WITHOUT UC CAMPUSES

	COST IMPS (K$/Year)	Line (K$/Year)	Total (M$/Year) (IMP Cost + Line Cost)
ARPA Base Network	605	825	1.43
Cost of Adding 7 U.C. Campuses to the Base Network	123	71	0.19
ARPA Base Network + 7 U.C. Campuses	728	896	1.62

	TRAFFIC KBITS/SEC/NODE	BITS/SEC/NODE PAIR
Among ARPA Base Network Nodes	11	550
Among 9 U.C. Campuses	3.2	400

Overall total throughput (ARPA + U.C.): 260 KBITS/SEC

Note:
1. The ARPA base network is shown in Figure 3.2.3(a).
2. The lines between ARPA base network nodes are all 50 KBITS/SEC.
3. The added lines are all 9.6 KBITS/SEC, except at node 25 which is inserted into the line linking node 2 and node 3.
4. The ARPA net including U.C. campuses is shown in Figure 3.2.6.

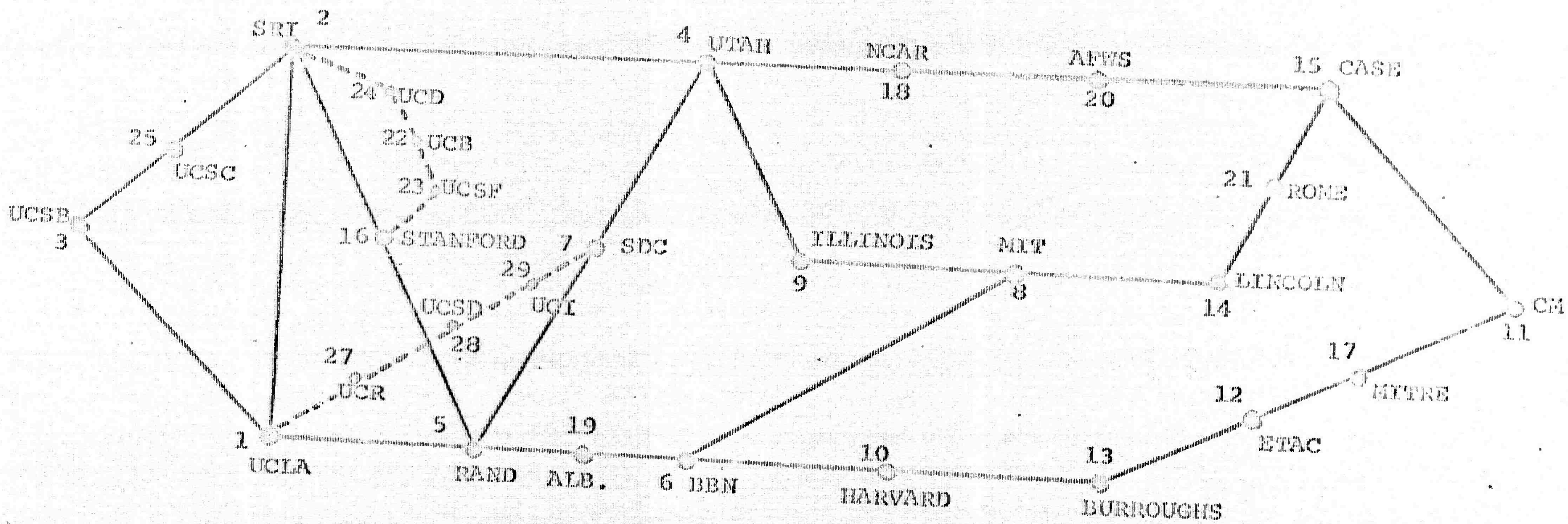
TABLE 3.2.7

UNIVERSITY OF CALIFORNIA'S NODE LOCATIONS

Node Number	Node Name	Node Location LAT.	LONG.
22	UC Berkeley	37 50	122 17
23	UC San Francisco	37 50	122 30
24	UC Davis	38 33	121 45
25	UC Santa Cruz	37	122
26	UC Irvine	33 35	117 50
27	UC Riverside	33 56	117 25
28	UC San Diego	32 55	117 20

All solid lines 50 kilobits/sec

All dotted lines 9.6 kilobits/sec



ARPA Network With U.C. Campuses

FIGURE 3.2.6

3.3. A PRELIMINARY COST ANALYSIS FOR HIGH THROUGHPUT NODES

Usage of the ARPA Network will differ from node to node. In general, there will be two types of users: those that require peak input capacity only moderately higher than their average capacity, and those who occasionally will require very high peak bandwidths in comparison to their average usage. This section represents an initial attempt to uncover the additional cost of serving a user in the second category.

Our object is to determine the cost of supplying a few nodes with the capability of inputting traffic at an 80 KB/Sec. rate while all other nodes are limited to a 10-15 KB/sec. rate. An initial average base traffic level of 10 KB/Sec/Node is assumed.

The method of cost computation operates as follows. The costs for 20, 40, 60, 80, and 100 node networks for the 10 KB/Sec/Node's traffic level are found. It is assumed that an arbitrary node wishes to transmit 80 KB/Sec. for short durations of time. If both the origin and the destination of this traffic is known, the best approach is to make a specific optimization run with this data to determine the cost of adding sufficient capacity to the network so that other users will not be affected.

To obtain guidelines for general cost analysis when both sender and receiver are unknown, we follow the steps below.

1. Compute the effective traffic load per node within the network when a single IMP is adding traffic at an 80 KB/Sec. rate to the net and all other IMPS are generating at a 10 KB/Sec. rate.

2. Determine the cost to construct a network to accommodate the effective traffic load with the specified delay constraint. This computation is based on the assumption that the load is uniformly distributed among all nodes. Let this cost be indicated by $\Delta GLOB$.

3. Determine the average cost of increasing the output capacity of a single IMP to 80 KB/Sec. from about 10-15 KB/Sec. This cost is assumed to be the cost of upgrading two average length lines from 10 KB/Sec. capacities to 50 KB/Sec. capacities. This cost is $\Delta LOC = 2\ [(850 - 650) + (5.00 - 0.40)$ Avg. Length] per month.

4. The average incremental cost to enable a subscriber to input traffic at the 80 KB/Sec. rate is then

$$\Delta \text{ Cost} = \Delta LOC + \Delta GLOB/NHT$$

where NHT is the number of nodes which desire high throughput capability.

<u>Note:</u> The equation for $\Delta COST$ assumes that either (1) high input rates occur infrequently or (2) only a few nodes will have

high injection rates. With either of these assumptions, it will be relatively unlikely that more than one user will be generating 80 KB/Sec. at the same time. Thus, the network need only be upgraded to handle one high input rate at a time, and the cost of this capacity increase can be shared by the NHT high throughput nodes. (Naturally, if specific requirements are available, more exact provisions can be made.)

Table 3.3.1 indicates relative cost factors and traffic loads for the systems considered. The last column in this table indicates the average yearly cost to supply exactly one node with high input capability. As the number of nodes requiring this capability increases, the cost per high throughput node decreases.

Figure 3.3.1 indicates cost per high throughput node as a function of the percentage of nodes with this capability. The curves for 20, 40, 60, 80, and 100 node networks have been computed, but it is interesting to note that the curves for 60, 80, and 100 node systems are virtually identical, and therefore only one curve has been drawn for these situations.

As a final point, we note that the cost-throughput curves for 20, 40, 60, 80, and 100 node networks are essentially linear in the 1 KB/Sec. - 15 KB/Sec. regions. This means that the incremental costs shown in the last three columns of Table 3.3.1

and illustrated in Figure 3.3.1 are independent of the base traffic load (which was initially assumed to be 10 mE/Sub/B.H.). This means that the incremental costs given are applicable for a range of traffic loadings.

TABLE 3.3.1

Number of Nodes	Cost for 10 KB/Sec/Node (K$/Year)	Equivalent Load for one 80 KB/Sec. Source (KB/Sec/Node)	Avg. Line Length (Miles)	ΔLOC (K$/Yr.)	ΔGLOB (K$/Yr.)	Total Cost to add for first high throughput node (K$/Year)
20	0.860	12.8K	300	$35K	$65K	$100K
40	1.605	11.4	250	30	108	138
60	2.448	10.95	250	30	78	108
80	3.850	10.71	225	27	110	137
100	4.655	10.57	225	27	149	176

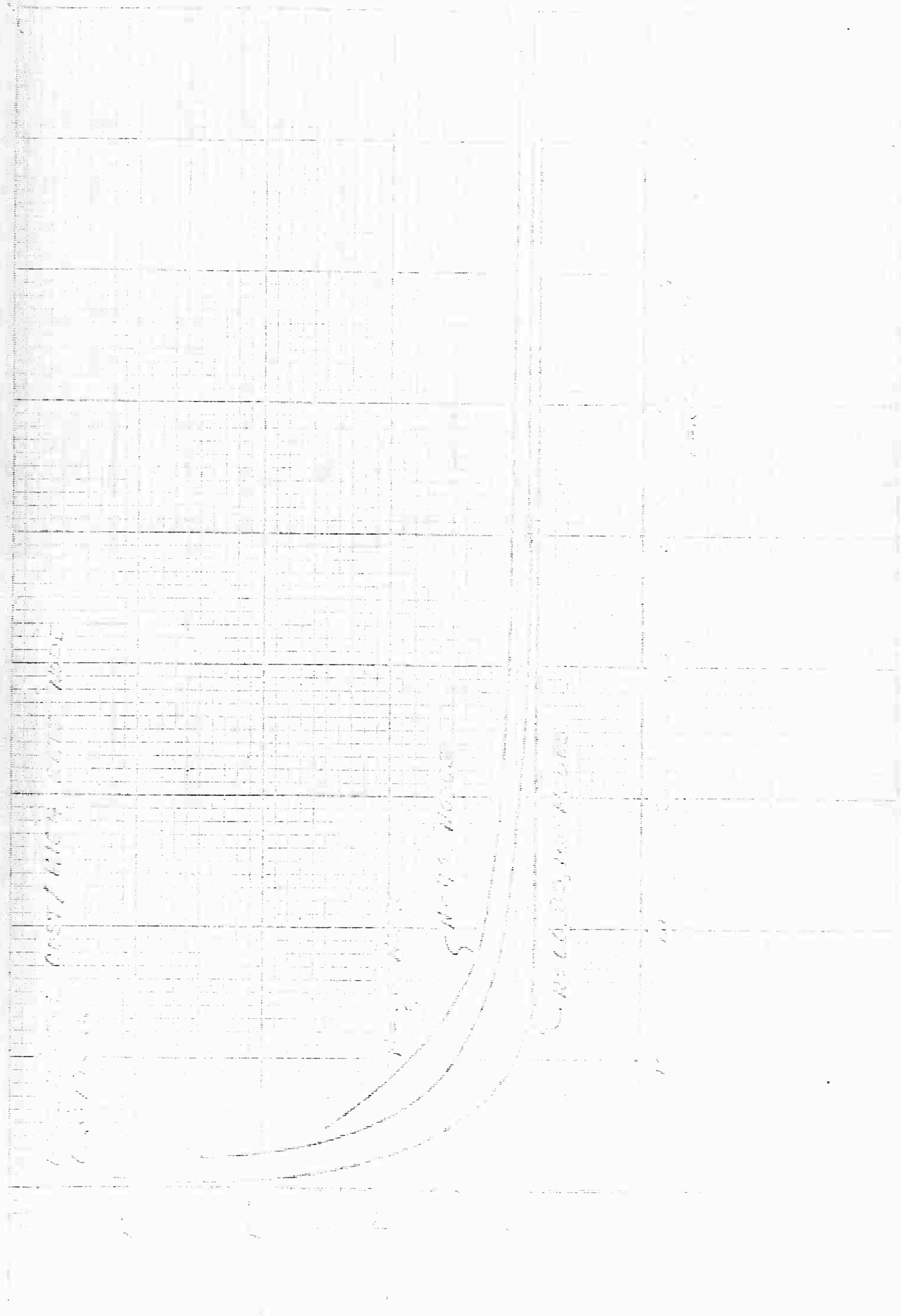


FIGURE B-2

3.4. A PRELIMINARY STUDY OF LOCAL DISTRIBUTION SCHEMES

As a computer-communication system grows, it is likely that more and more nodes will fall into the same metropolitan area. In such cases, it is often more convenient from the standpoint of installation and maintenance to restrict the structure of the local networks. This study is to investigate different design concepts for connecting nodes in the same metropolitan area to a nationwide network. Three concepts are considered:

1. Generalized Star Structures. Every local node is connected directly to one or two DDP 516's which are located at telephone company central offices. The central DDP 516's are connected to remote nodes outside the region. In this way the DDP 516's sole function is message store-and-forward, and the local nodes do not require any message switching functions.

2. Locally-looped Structures. Several local nodes are first connected into a chain. The two ends of the chain are then connected to two DDP 516's located in telephone company central offices. In this case, the local nodes perform a limited amount of message switching.

3. Unconstrained Structures. No intermediate DDP 516's at central offices are used. The local nodes are first connected together. Some of them are then connected to nodes outside the local metropolitan area.

The most economical design depends on node location, traffic patterns and other criteria. Lacking this specific information, case studies cannot be performed. However, we can set up and evaluate an abstract model to weigh the trade-offs between the three design concepts. The results obtained from studying this model, we believe, can be applied to most real problems.

The model is described as follows:

1. The metropolitan area is a 50-mile by 50-mile square.

2. The square is divided into a 5-mile spaced grid. The local nodes and/or the DDP 516's can be located only at the intersecting points. We make this assumption partly because telephone company central offices are distributed quite uniformly in a metropolitan area, and because in many cases the line charge between two points is based on the distance between the two corresponding central offices closest to them.

3. The node locations are then generated randomly.

4. The network topology and traffic pattern are based somewhat on the New York region in a 200-node case studied in Section 3.5 in which Los Angeles, San Jose, Jersey City and Hartford are directly connected to the New York metropolis.

Two designs are picked for the centralized star network in our study. One is chosen to minimize the sum of total length

of the lines between the DDP 516's, and the lines from the DDP 516's to distant nodes. (These lines are usually more expensive per unit length than other local lines.) The other configuration is chosen to minimize a heuristic based on the total length of all lines.

Using the through traffic as a parameter, we vary the level of the locally-originated traffic. The local network annual costs are obtained for various local traffic levels and are plotted in Figures 3.4.1 and 3.4.2 for a 10-node local and in Figures 3.4.3 and 3.4.4 for a 20-node local network.

From the curves we can draw the following observations:

1. The fewer the DDP 516's at telephone company central offices, the lower the cost.

2. With reasonable traffic levels and with proper network topology, the effects of the through traffic and the locally-originated traffic are almost independent of one another.

3. When the through traffic does affect the total local traffic which can be generated, either (1) local cost is slightly higher, or (2) locally-originated traffic slightly lower for higher through traffic.

The generalized star design evaluated in this study is one-connected. As expected, it is not as reliable in most cases as the unconstrained design. The one-connected star structure is already more expensive to build than the constrained design. To make the former design two-connected and thus more reliable, the cost would have to be significantly higher. Reliability analyses for both the 10-node and 20-node local networks are listed in Table 3.4.1.

TABLE 3.4.1

	Reliability		Percentage of Unconnected Node Pairs	
	Link	Node	Star Structure	Unconstrained Structure
10 nodes	.98	1.00	3.9	2.0
	.96	1.00	7.7	4.9
	.94	1.00	11.0	7.7
	.98	0.9932	6.2	1.6
	.96	0.9868	11.6	9.0
	.94	0.98	16.8	14.2
20 nodes	.98	1.00	3.6	2.0
	.96	1.00	7.6	6.1
	.94	1.00	11.1	11.0
	.98	0.9932	6.1	4.3
	.96	0.9868	12.3	11.2
	.94	0.98	16.4	18.9

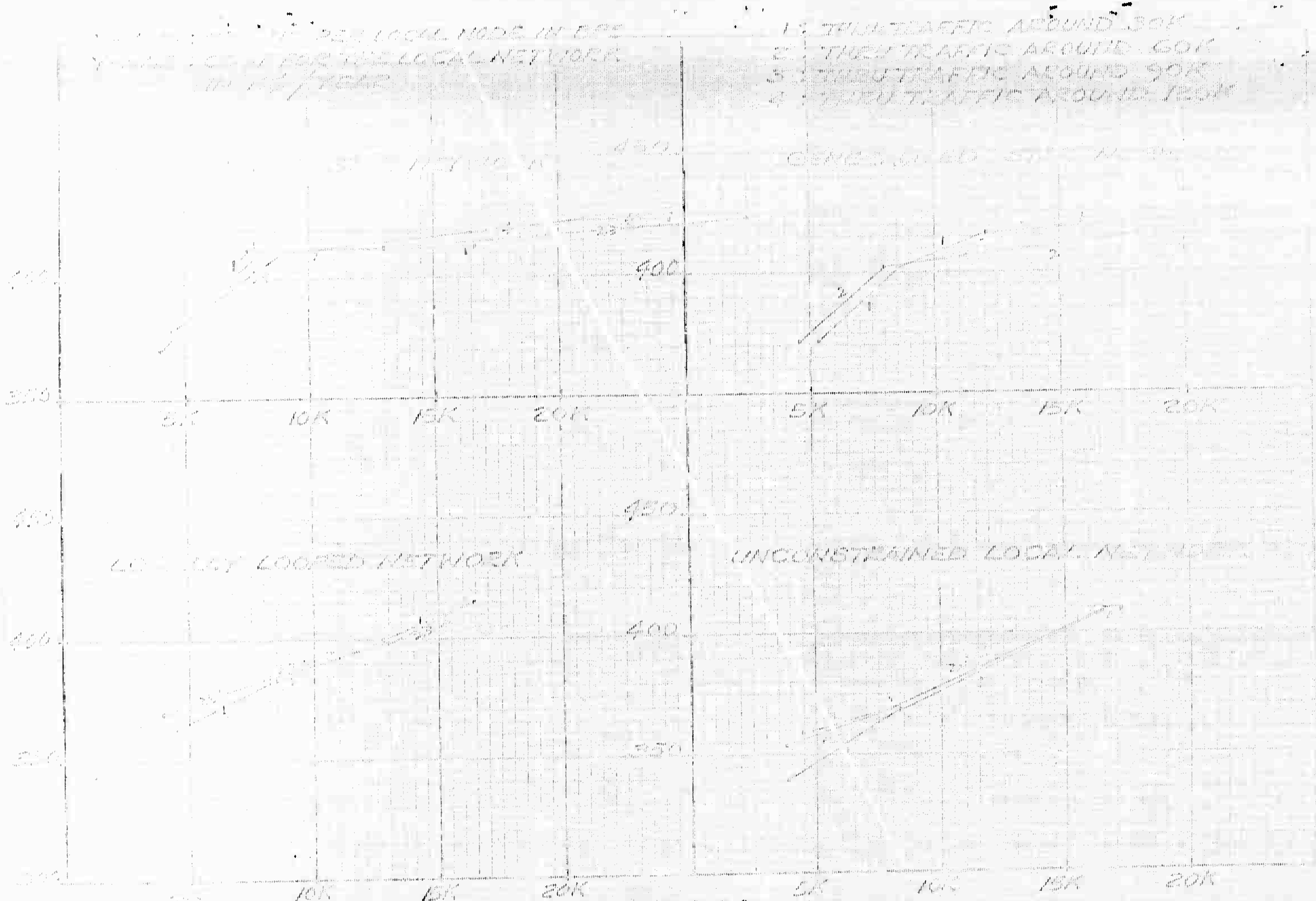


FIGURE 3.4.1

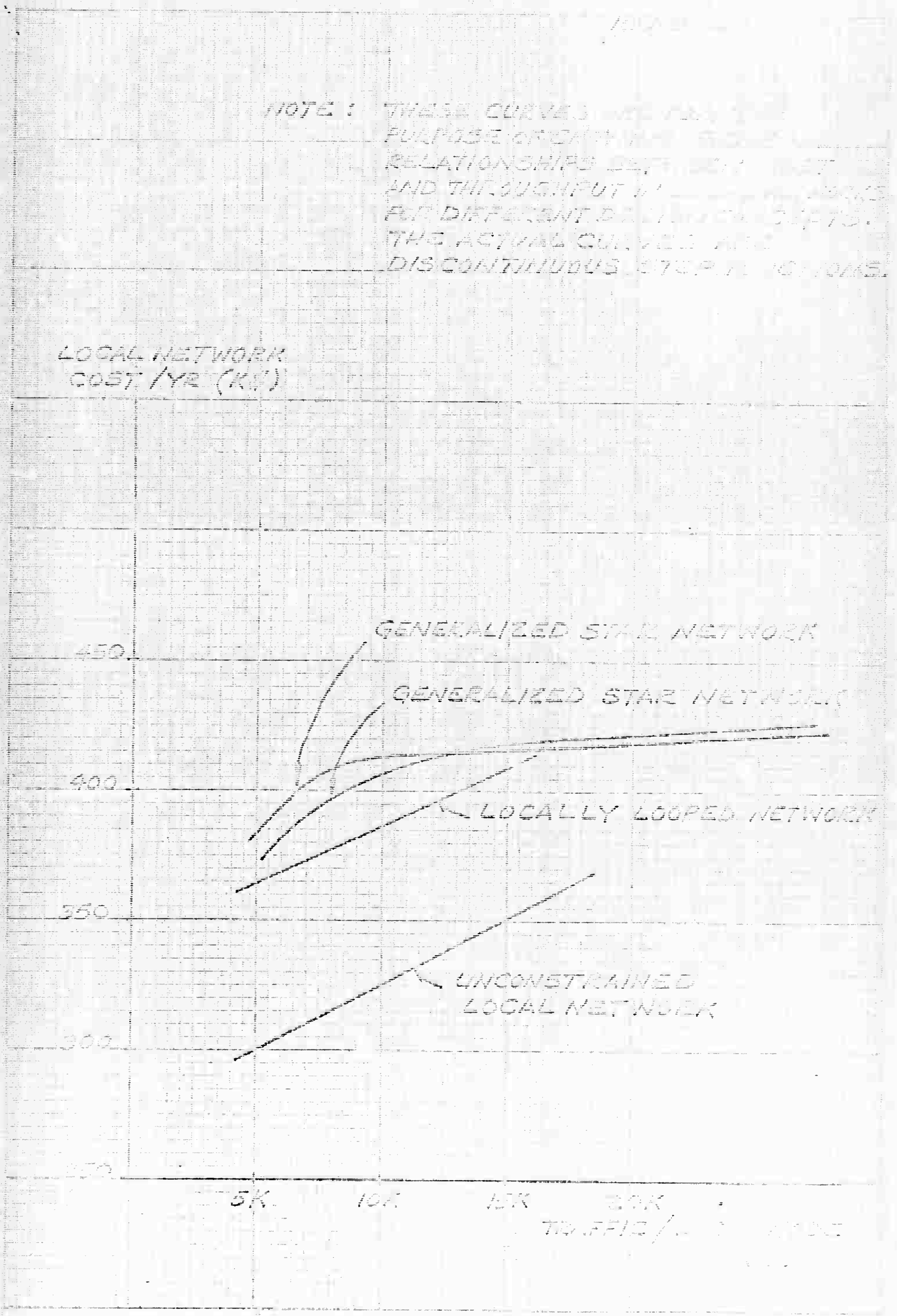

NOTE :
LOCAL NETWORK
GENERALIZED STAR NETWORK
GENERALIZED STAR NETWORK
LOCALLY LOOPED NETWORK
UNCONSTRAINED
LOCAL NETWORK
450
400
350
300
5K
10K

FIGURE 3.4.2

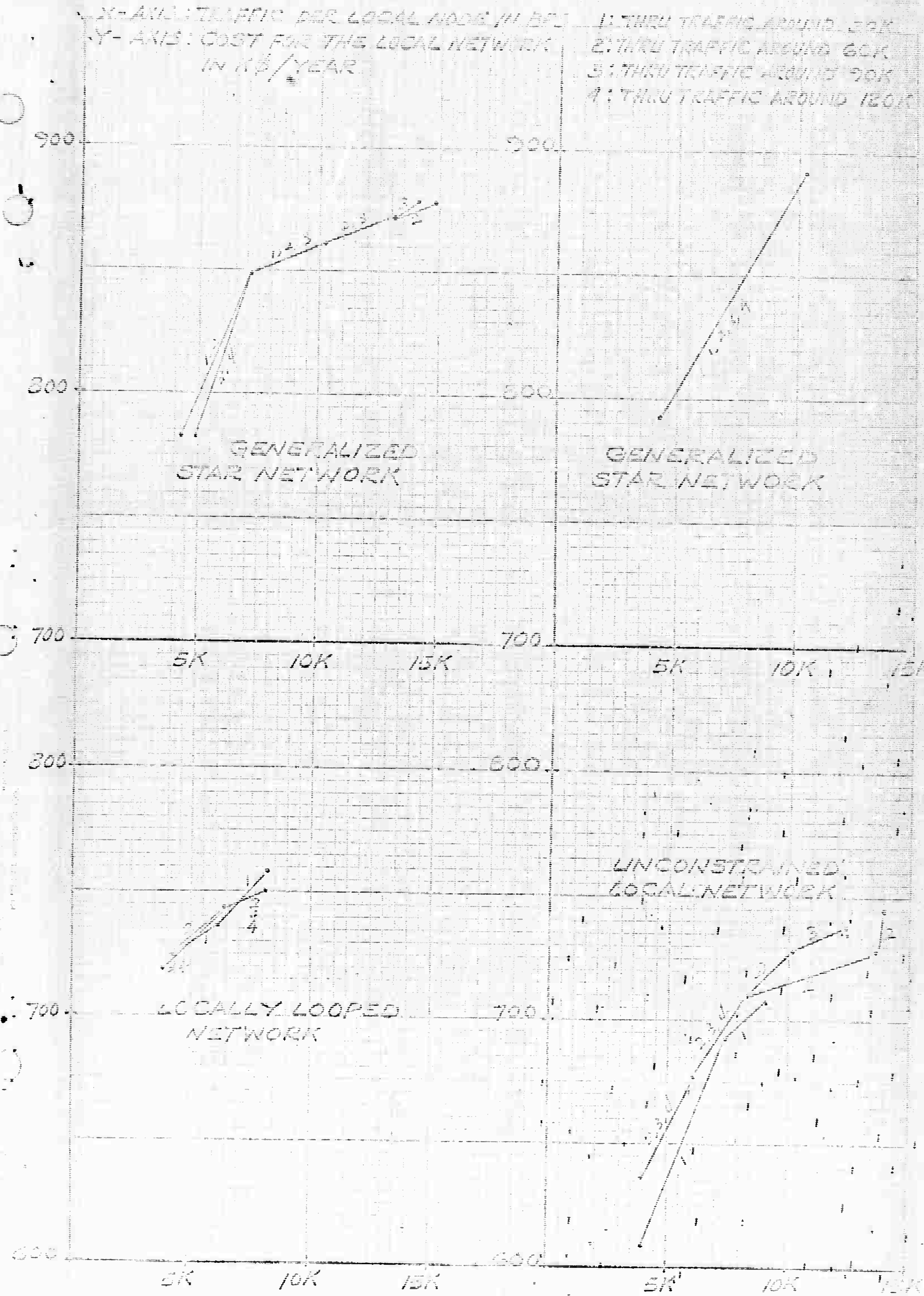

X-AXIS: TRAFFIC PER LOCAL NODE IN BPS
Y-AXIS: COST FOR THE LOCAL NETWORK IN K$/YEAR
1: THRU TRAFFIC AROUND 30K
2: THRU TRAFFIC AROUND 60K
3: THRU TRAFFIC AROUND 90K
4: THRU TRAFFIC AROUND 120K
900
800
700
5K
10K
15K
GENERALIZED STAR NETWORK
GENERALIZED STAR NETWORK
800
700
600
LOCALLY LOOPED NETWORK
UNCONSTRAINED LOCAL NETWORK

FIGURE 3.4.3

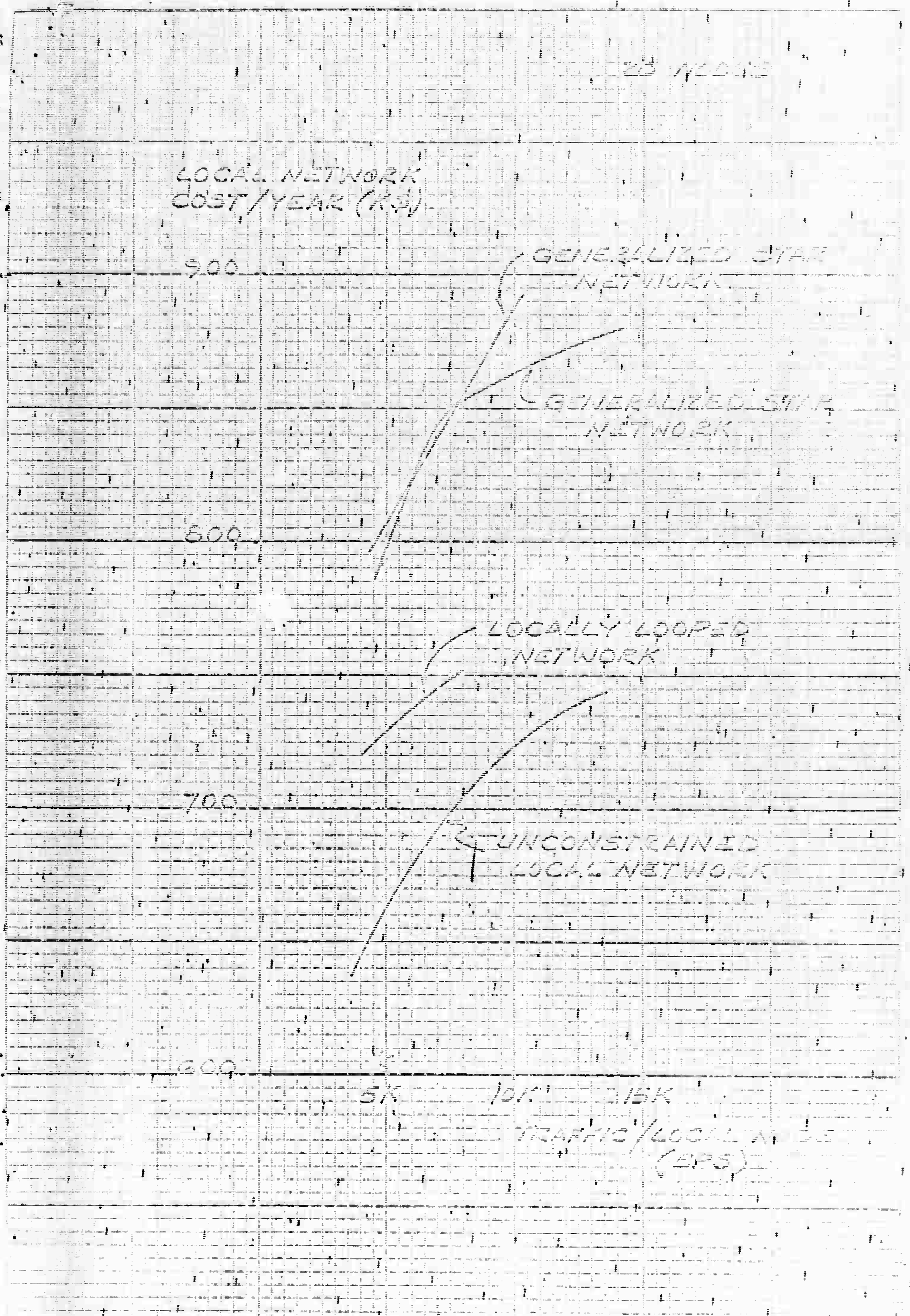
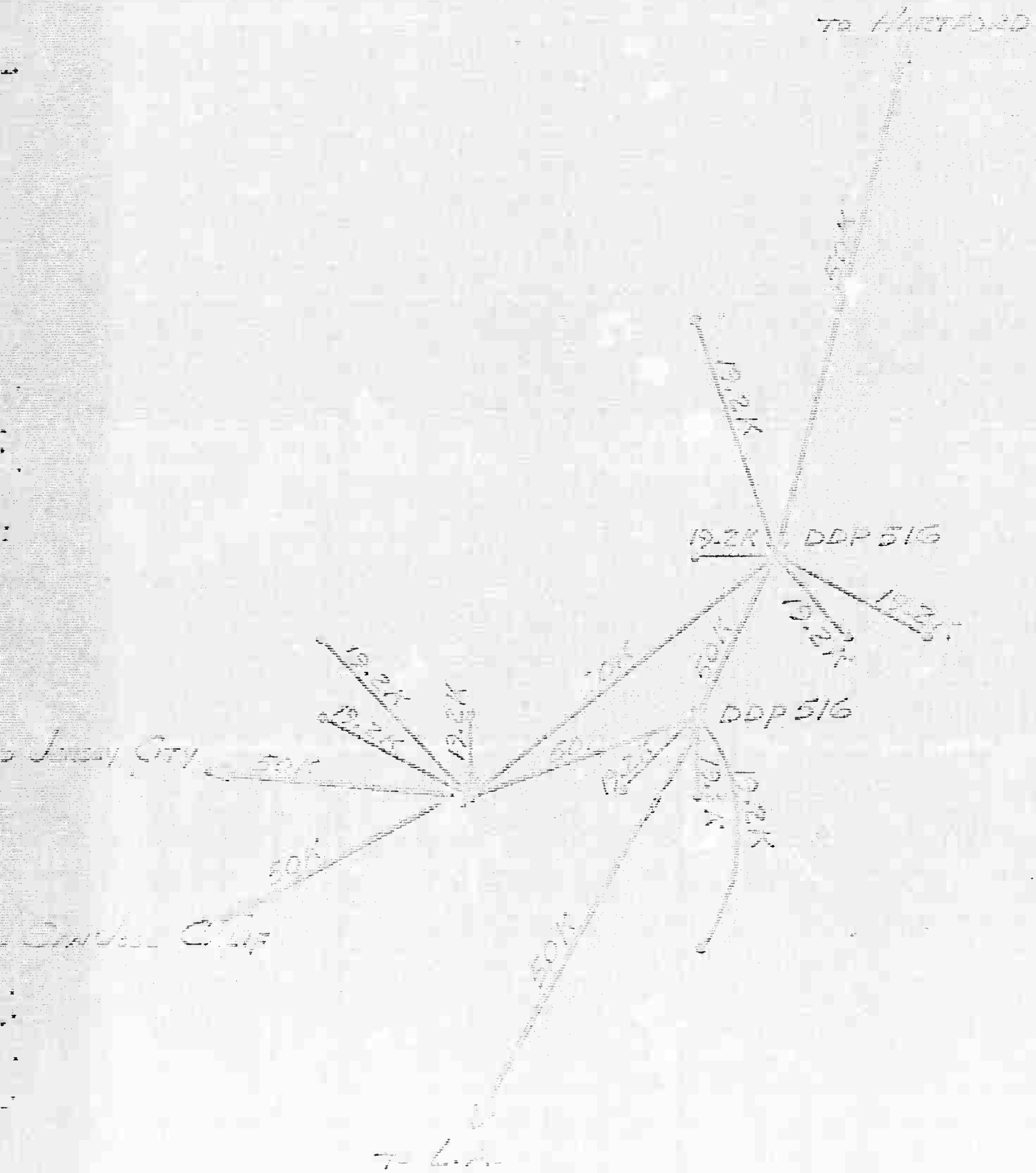
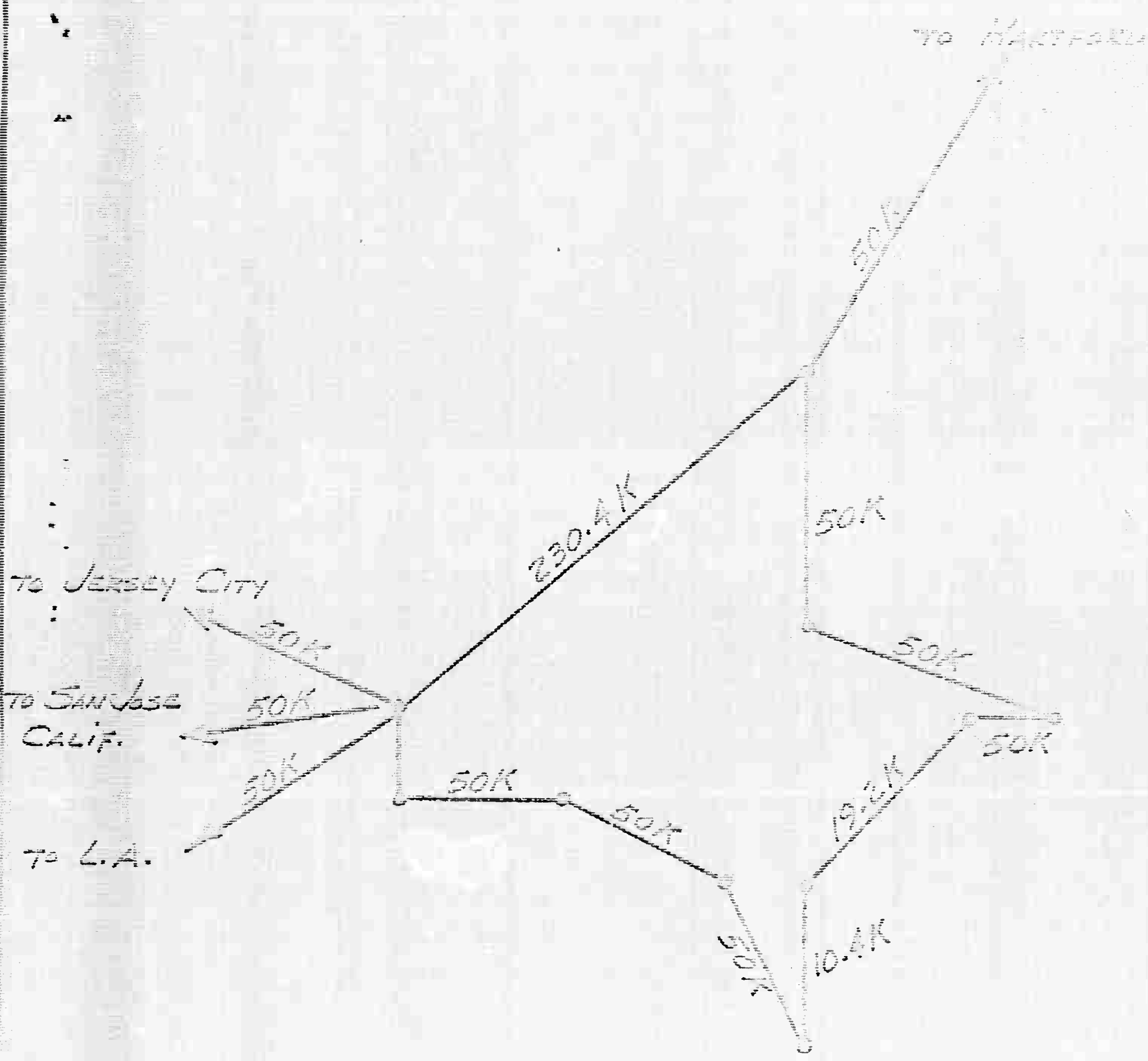

28 NODES
LOCAL NETWORK COST/YEAR (K$)
900
800
700
600
GENERALIZED STAR NETWORK
GENERALIZED STAR NETWORK
LOCALLY LOOPED NETWORK
UNCONSTRAINED LOCAL NETWORK
5K
10K
15K
TRAFFIC/LOCAL NODE (BPS)

FIGURE 3.4.4



[illegible] DESIGN
FOR 9KBPS / NODE TRAFFIC

FIGURE 3.4.5



10 NODE UNCONSTRAINED DESIGN
FOR 6 KBPS/NODE TRAFFIC

FIGURE 3.4.6

3.5. A 200 NODE STORE-AND-FORWARD NETWORK

In this section we describe two 200-node networks that were designed as part of our study of cost-throughput characteristics of large store-and-forward systems. The networks were designed to accommodate input rates of 3.1 and 8.0 kilobits per second per node. Thus, the 8 kilobit system indicates an upper bound on the cost of the ARPA Network if it expands to a 200-node size.

First, we summarize the factors which influence the network designs:

1. The system considered contains 200 IMPs located in the 62 largest metropolitan areas in the Continental United States.

2. Required traffic between any two IMPs is independent of distance. From an IMP in city C_i with population P_i to an IMP in city C_j with population P_j, the traffic flow requirement is

$$\frac{K(P_i/[P_i/R])(P_j/[P_j/R])}{\sum_k (P_k/[P_k/R])}$$

where K is a positive constant, R is the required population per IMP, and [x] is the smallest integer no greater than x.

3. Messages are assumed to have the same packet structure and formats as in the ARPA Network as described in Reference [4]. 65% of all messages are assumed to be single packets.

4. In any acceptable network design, a minimum of two nodes and/or links must fail before all paths are broken between any pair of nodes.

5. Throughput is equal to the average number of bits/second/node which causes an average short message response time of 0.5 seconds.

6. Only hardware presently being used in the ARPA Network is used in any design. Only communication link options presently available are used in the design. Cost factors used are shown in Tables 3.5.1 and 3.5.2.

TABLE 3.5.1

LINE COSTS

Capacity	Fixed Cost/Month	Cost Per Mile/Month
10,400 bps	$650.00	$0.40
19,200 bps	$850.00	$2.50
50,000 bps	$850.00	$5.00
230,400 bps	$1,300.00	$30.00

All lines fully duplex

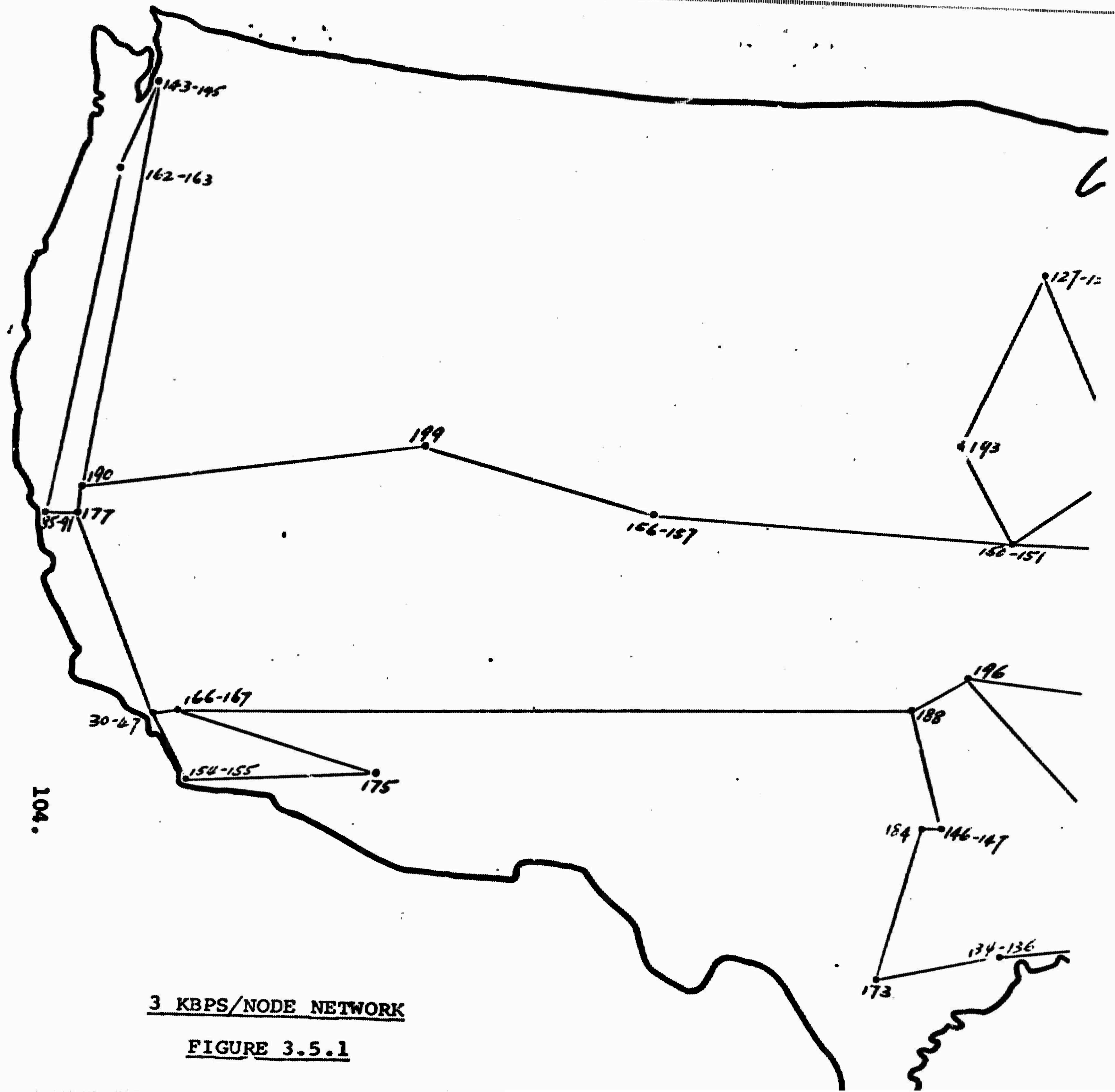
TABLE 3.5.2

NODE COSTS

Description	Rental Cost Per Year	Maintenance Cost Per Year
DDP-516 IMP with up to 7 fully duplex I/O channels	$25,700	$7,600
DDP-316 IMP with up to 5 fully duplex I/O channels. Processing rate is 3/4 that of 516 IMP.	$12,600	$5,000

Figures 3.5.1 and 3.5.2 show simplified versions of the 3 and 8 KBPS/node networks while their characteristics are indicated in Tables 3.5.3 and 3.5.4.

Figure 3.5.3 integrates the 200 node data with the 20, 40, 60, 80, and 100 node data described in NAC's Second Semi-Annual Technical Report. In this figure each point represents a network. The number next to a point is its throughput. It is easy to see that the 200 node points verify the trends derived from the 20, 40, 60, 80, and 100 node data. This establishes the soundness of the ARPA Network technology for the systems considered.



3 KBPS/NODE NETWORK

FIGURE 3.5.1

TABLE 3.5.3

SUMMARY OF OVERALL NETWORK DESIGN AND PERFORMANCE

COST/YEAR

TIPS	\$3520000
IMPS	\$3424054
TOTAL	\$6944054
PER NODE	\$34 20 AVERAGE
PER PACKETS	35.49¢ AVERAGE, BASED ON 24 HRS/DAY OPERATION

H.C. TRAFFIC

PER NODE IN CLOSE NET	10.7 KILOBITS/SECOND AVERAGE
PER NODE IN OVERALL NET	3.1 KILOBITS/SECOND AVERAGE
TOTAL	621 KILOBITS/SECOND AVERAGE

IMPS

DDP 316		DDP 516	
DEGREE	NUMBER	DEGREE	NUMBER
1	0	4	0
2	168	5	0
3	32	6	0
Total	200	Total	0

TOTAL NUMBER OF IMPS = 200

AVERAGE DEGREE = 2.16

LINKS

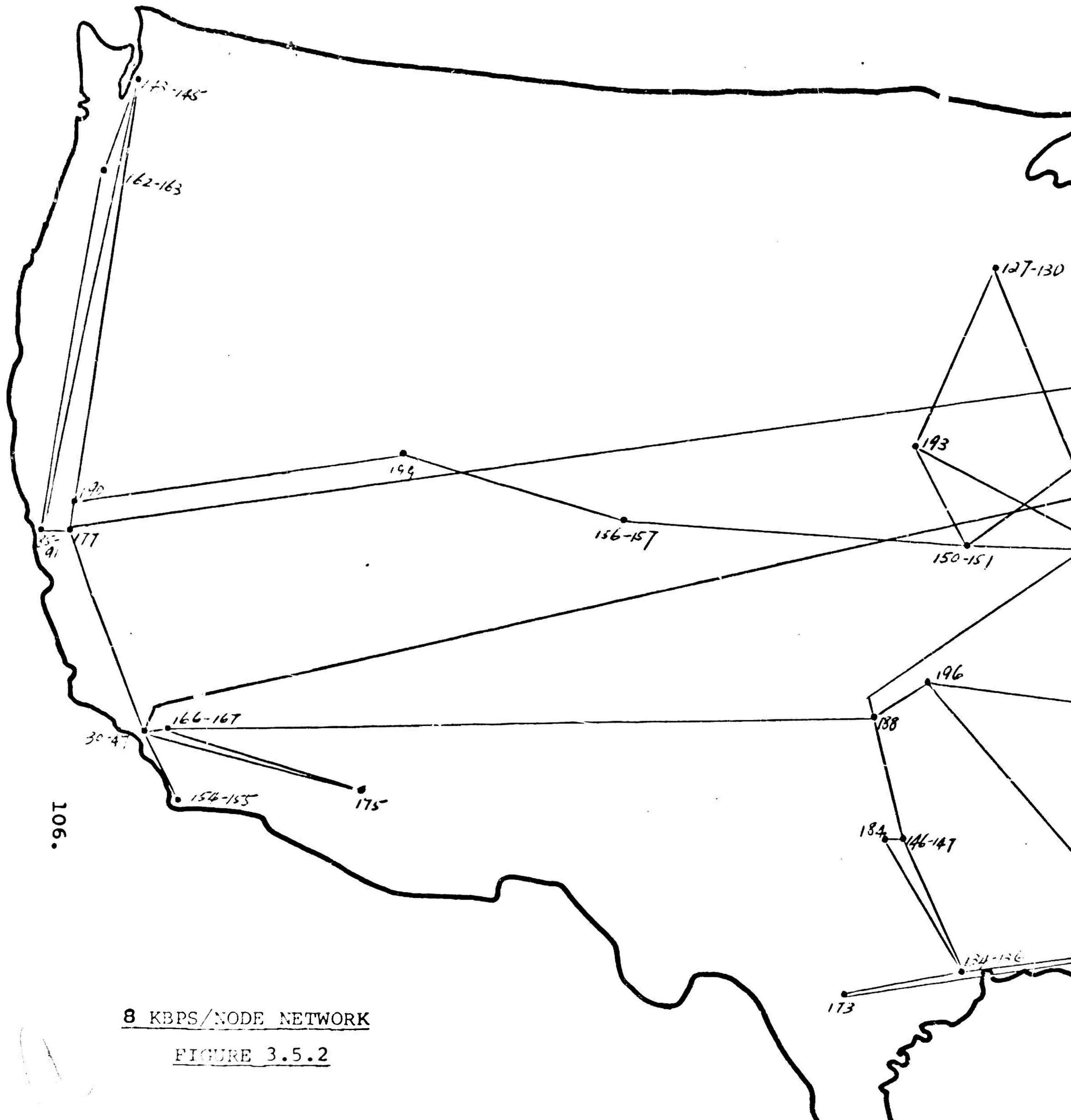
CAPACITY	NUMBER	LENGTH
10400	105	6181 MILES
19200	66	5141 MILES
50000	10	423 MILES
100000	25	6853 MILES
230400	10	511 MILES
TOTAL	216	19098 MILES

AVERAGE LINK LENGTH = 88 MILES

NUMBER OF METROPOLITAN AREAS = 62

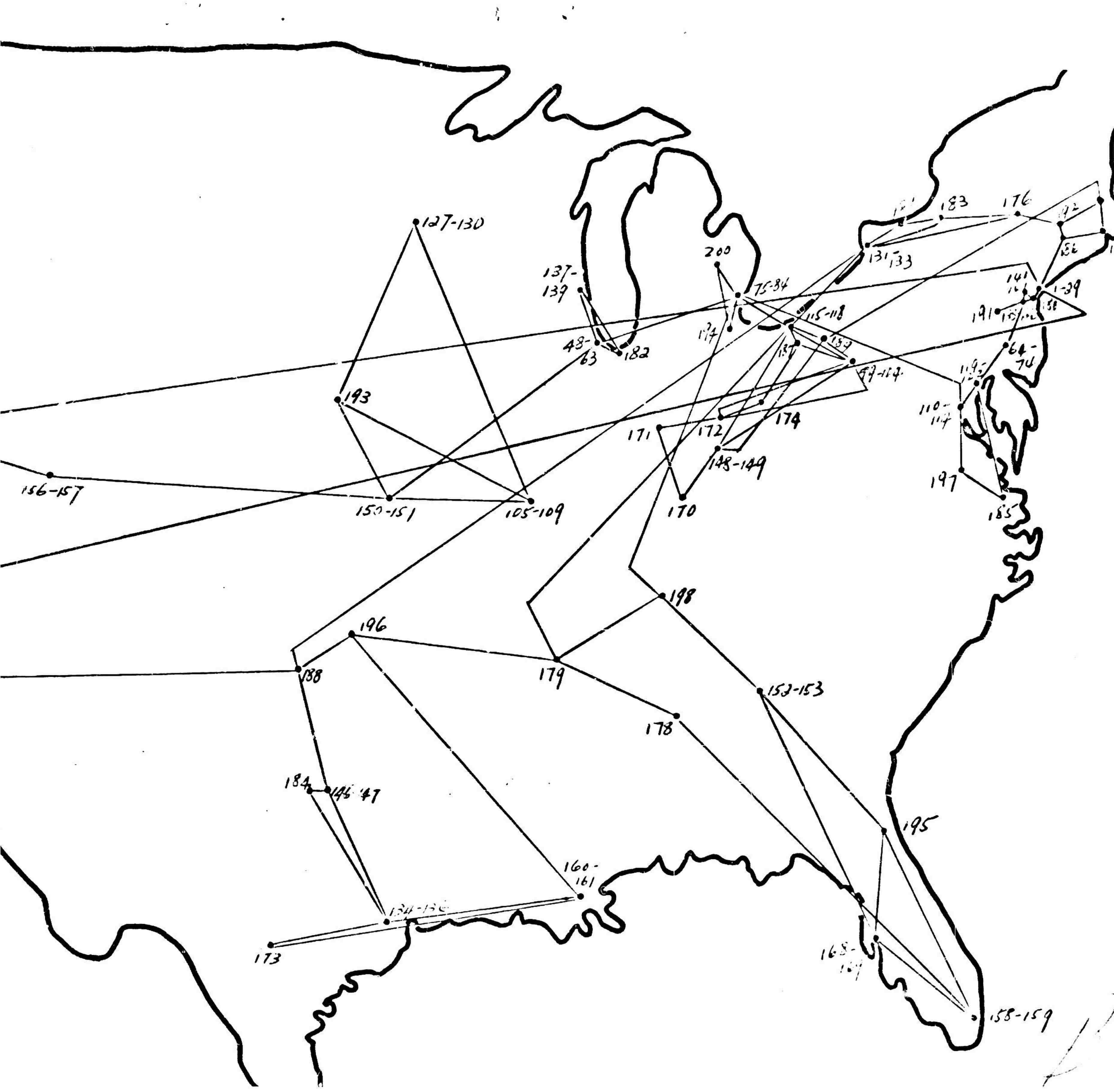
AVERAGE HOST-HOST SMALL MESSAGE TIME DELAY = .344 SECONDS

BLANK PAGE



8 KBPS/NODE NETWORK

FIGURE 3.5.2



127-130
137-
139
48-
63
182
200
75-84
115-118
193
156-157
150-151
105-109
171
172
174
148-149
170
183
176
131-
133
141
1-39
191
64-
70
110-
114
197
185
198
196
179
188
178
152-153
184
195
160-
161
173
168-
158-159

TABLE 3.5.4

SUMMARY OF OVERALL NETWORK DESIGN AND PERFORMANCE

COSTS		MAX TRAFFIC	
IMPS	\$3677000	PER NODE IN GLOBE NET	23.2 KILOBITS/SECOND AVERAGE
LINES	\$5405028	PER NODE IN OVERALL NET	8.0 KILOBITS/SECOND AVERAGE
TOTAL	\$9082028	TOTAL	1600 KILOBITS/SECOND AVERAGE
PER NODE	\$ 45410 AVERAGE		
PER MEGABITS	18.00 CENTS AVERAGE, BASED ON 24 HRS/DAY OPERATION		

IMPS

DDP 316		DDP 516	
DEGREE	NUMBER	DEGREE	NUMBER
1	0	4	10
2	100	5	0
3	90	6	0
TOTAL	190	TOTAL	10

TOTAL NUMBER OF IMPS = 200
AVERAGE [illegible] = 2.55

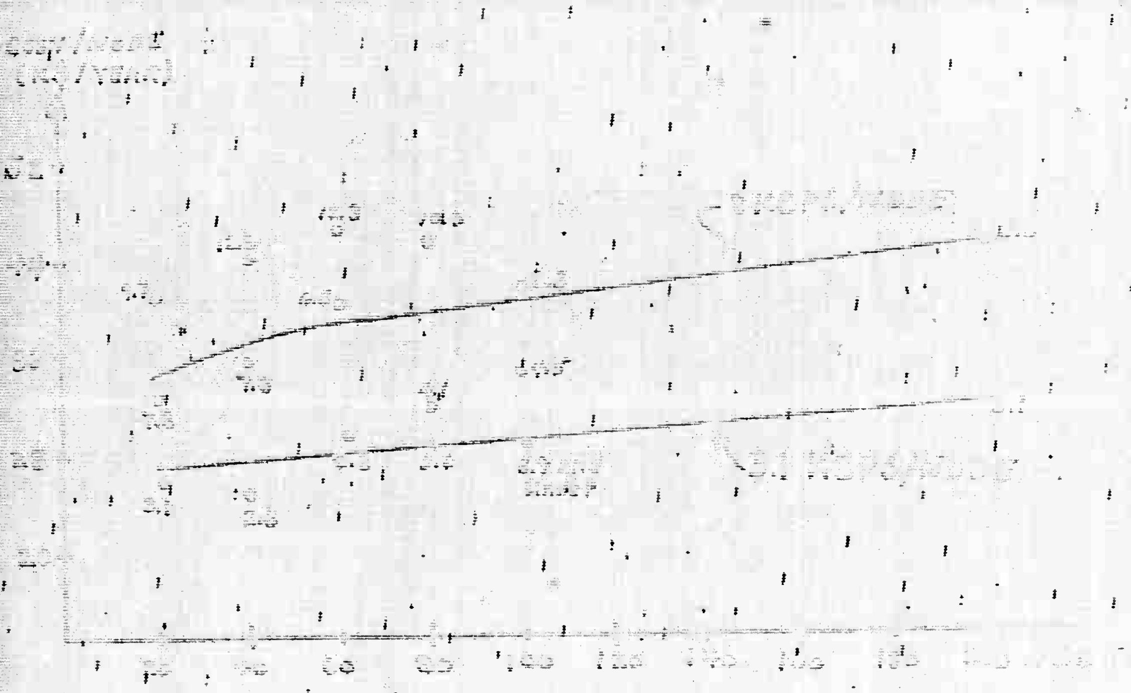
LINES

CAPACITY	NUMBER	LENGTH
10400	99	8439 MILES
19200	50	3988 MILES
50000	61	8738 MILES
100000	21	10508 MILES
230400	24	2222 MILES
TOTAL	255	33896 MILES

AVERAGE LINK LENGTH = 133 MILES

NUMBER OF METROPOLITAN AREAS = 62

AVERAGE HOST-HOST SMALL MESSAGE TIME DELAY = .270 SECONDS



COSTS AND THROUGHPUT FOR
20, 40, 60, 80, 100 and
200 NODE NETWORKS

FIGURE [illegible]

REFERENCES

1) Berge, C., The Theory of Graphs and Its Applications, Wiley, [1962].

2) Fisz, M., Probability Theory and Mathematical Statistics, Wiley [1963].

3) Frank, H. and Frisch, I.T., "Analysis and Design of Survivable Networks", IEEE Transactions on Communications Technology, Vol COM-18 No. 5, [October 1970].

4) Frank, H., Frisch, I.T., and Chou, W., "Topological Considerations in the Design of the ARPA Computer Network," AFIPS Conference Proceedings, Vol. 36, AFIPS Press, pp. 581-587, [1970].

5) Fulkerson, D. R. "Networks, Frames, Blocking Systems," Mathematics of the Decision Sciences, Part I, pp. 303-334, American Mathematical Society, [1968].

6) Gomory, R. E. and Hu, T C., "Multi-Terminal Network Flows," J. Siam, Vol. 9, No. 4 pp. 551-570 [1961].

7) Jacobs, I.M., Connectivity in Probabilistic Graphs - Technical Report 356, Massachusetts Institute of Technology, Research Laboratory of Electronics, Cambridge, Mass. [1959].

8) Kruskal, J. B., "The Number of Simplices in a Complex," in Mathematical Optimization Techniques, ed. R. Bellman, Berkeley: University of California Press 251-278 [1963].

9) [illegible], J. D., Synthesis of Reliable Communication Networks PhD dissertation, University of Pennsylvania [1968].

10) Moore, E. F. and C. E. Shannon, "Reliable Circuits Using Less Reliable Relays," J. of the Franklin Institute, Vol. 262, Pt.I, pp. 191-208, [1956].

11) Seshu, S. and Reed, M.B. Linear Graphs and Electrical Networks, Addison-Wesley [1971].